

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: February 17, 2004, 10:50:12 ; Search time 17.1089 Seconds  
(without alignments)  
83.497 Million cell updates/sec

Title: US-09-900-147-2

Perfect score: 42

Sequence: 1 NVLMAMNII 9

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1107863 seqs, 150726573 residues

Total number of hits satisfying chosen parameters: 717921

Minimum DB seq length: 0

Maximum DB seq length: 100

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : A Geneseq\_19Jun03.\*

1: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1980.DAT.\*  
2: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1981.DAT.\*  
3: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1982.DAT.\*  
4: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1983.DAT.\*  
5: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1984.DAT.\*  
6: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1985.DAT.\*  
7: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1986.DAT.\*  
8: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1987.DAT.\*  
9: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1988.DAT.\*  
10: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1989.DAT.\*  
11: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1990.DAT.\*  
12: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1991.DAT.\*  
13: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1992.DAT.\*  
14: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1993.DAT.\*  
15: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1994.DAT.\*  
16: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1995.DAT.\*  
17: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1996.DAT.\*  
18: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1997.DAT.\*  
19: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1998.DAT.\*  
20: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1999.DAT.\*  
21: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA2000.DAT.\*  
22: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA2001.DAT.\*  
23: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA2002.DAT.\*  
24: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA2003.DAT.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	42	100.0	9	AAW30502	DP-1 transcription
2	42	100.0	16	AAW30506	DP-1 transcription
3	42	100.0	19	AAW30516	DP-1 transcription
4	42	100.0	19	AAW30504	DP-1 transcription
5	42	100.0	20	AAW30505	DP-1 transcription
6	42	100.0	28	AAW57051	E2F activity inhib
7	42	100.0	28	AAW57055	E2F activity inhib
8	42	100.0	30	AAW30507	DP-1 transcription
9	42	100.0	37	AAW30501	DP-1 transcription

10	38	90.5	19	AAW30515	DP-1 transcription
11	37	88.1	83	AAW32163	Soybean DP-1 prote
12	36	85.7	19	AAW30517	DP-1 transcription
13	34	81.0	14	AAW30511	DP-1 transcription
14	31	73.8	67	AAW58362	Human brain expres
15	29	69.0	8	AAU72578	Plant dimerisation
16	29	69.0	88	AAW77753	Pyruvate oxidase p
17	29	69.0	89	ABP38852	Staphylococcus epi
18	29	69.0	100	ABP38589	Staphylococcus epi
19	28	66.7	15	AAW57052	E2F activity inhib
20	28	66.7	56	AAW67766	Fragment from a wh
21	28	66.7	91	AAO00494	Human polypeptide
22	27	64.3	83	AAU44499	Propionibacterium
23	26	61.9	48	AAU21311	Human novel foetal
24	26	61.9	55	AAU75292	Human colon cancer
25	26	61.9	58	ABP03723	Human ORFX protein
26	26	61.9	62	ABP17490	Human nervous syst
27	26	61.9	71	AAO11790	Human polypeptide
28	26	61.9	72	ABP03425	Human ORFX protein
29	26	61.9	79	AAW20232	H. pylori cytoplas
30	26	61.9	96	AAV10972	H. pylori ORF 03ee
31	25	59.5	14	AAW00188	Human protein frag
32	25	59.5	25	AAW38530	Peptide #12567 enc
33	25	59.5	32	AAW63496	Rat REC20 serotoni
34	25	59.5	35	AAW76229	Fragment of human
35	25	59.5	51	ABG47854	Human liver peptid
36	25	59.5	51	ABB27831	Human peptide #482
37	25	59.5	51	ABB33002	Peptide #508 encod
38	25	59.5	51	ABB18475	Protein #474 encod
39	25	59.5	51	AAW53800	Human brain expres
40	25	59.5	51	AAW66186	Human bone marrow
41	25	59.5	51	AAW14057	Peptide #491 encod
42	25	59.5	51	AAW26463	Peptide #500 encod
43	25	59.5	51	AAW01797	Peptide #479 encod
44	25	59.5	51	ABG35836	Human peptide enco
45	25	59.5	52	AAU14232	Human novel protei

#### ALIGNMENTS

#### RESULT 1

AAW30502

ID AAW30502 standard; Peptide; 9 AA.

XX AC AAW30502;

XX DT 26-OCT-1998 (first entry)

XX DE DP-1 transcription factor antagonist peptide motif.

XX DE DP-1; transcription factor; antagonist; E2F protein; apoptosis;

XX DE cell proliferation; cardiovascular cell; restenosis; tumour;

XX DE surgical stent; therapy.

XX OS Synthetic.

XX OS Homo sapiens.

XX OS WO9828334-A1.

XX PD 02-JUL-1998.

XX PF 22-DEC-1997; 97WO-GB03506.

XX PR 20-DEC-1996; 96GB-0026589.

XX PA (PROL-) PROLIFIX LTD.

XX PI Bandara LR, La Thangue NB;

XX PR WPI; 1998-377596/32.

XX PT Polypeptide fragments of the DP-1 transcription factor - used for

PT inducing apoptosis, specifically in tumour and cardiovascular cells,  
 XX e.g. for preventing restenosis  
 XX  
 PS Claim 3; Page 44; 55pp; English.  
 XX  
 CC This peptide comprises amino acid residues 175-183 in the DEF box  
 CC region (see AAW30501) of transcription factor DPl. Claimed peptides  
 CC (II) (see AAW30502-07) containing this and/or another motif (see  
 CC AAW30503) of the DEF box, are antagonists of the heterodimerisation  
 CC of a DP protein with an E2F protein. Also claimed are variants  
 CC of these peptides, especially containing substitutions of residues  
 CC corresponding to residues 167, 169, 171 and 175 of DP-1, fusion  
 CC proteins (III) comprising (I) or (II) and a membrane translocation  
 CC sequence (see AAW30508), expression vectors encoding (I)-(III) and  
 CC host cells. (I)-(III) are used therapeutically to induce apoptosis,  
 CC specifically in tumour or cardiovascular cells, either in vivo or in  
 CC vitro, e.g. for purging bone marrow. Surgical stents comprising  
 CC (I)-(III) are used to treat or prevent restenosis in patients who  
 CC have undergone angioplasty. (I)-(III) function by inactivating  
 CC the DNA-binding activity of DP/E2F heterodimers. They are also  
 CC used as research reagents, as positive controls in assays for  
 CC identifying antagonists of DP-1/E2F dimerisation and as immunoassay  
 CC agents. Also described is the use of sequences antisense to  
 CC nucleic acids encoding (I)-(III) to control DP levels in cells,  
 CC particularly by gene therapy. When formulated with cytotoxic  
 CC or cytostatic agents, (I)-(III) enhance cell killing.  
 XX  
 SQ Sequence 9 AA;

Query Match 100.0%; Score 42; DB 19; Length 9;  
 Best Local Similarity 100.0%; Pred. No. 9.3e+05;  
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 NVLMAMNII 9  
 DB 1 NVLMAMNII 9

RESULT 2  
 AAW30506  
 ID AAW30506 standard; Peptide; 16 AA.  
 XX  
 AC AAW30506;

26-OCT-1998 (first entry)  
 DT  
 DE DP-1 transcription factor antagonist peptide H5.

XX DP-1; transcription factor; antagonist; E2F protein; apoptosis;  
 KW cell proliferation; cardiovascular cell; restenosis; tumour;  
 KW surgical stent; therapy.  
 XX

OS Synthetic.  
 OS Homo sapiens.

XX Key Location/Qualifiers  
 FH Peptide 5..11  
 FT /note= "Claim 3"  
 FT Peptide 7..16  
 FT /note= "Claim 3"

XX WO9828334-A1.

XX 02-JUL-1998.

XX 22-DEC-1997; 97WO-GB03506.

XX 20-DEC-1996; 96GB-0026589.

XX (PROL-) PROLIFIX LTD.

XX Bandara LR, La Thangue NB;

XX

DR WPI; 1998-377596/32.  
 XX  
 PT Polypeptide fragments of the DP-1 transcription factor - used for  
 PT inducing apoptosis, specifically in tumour and cardiovascular cells,  
 PT e.g. for preventing restenosis  
 XX  
 PS Claim 4; Page 44; 55pp; English.

CC Peptide H5 comprises amino acid residues 168-183 in the DEF box  
 CC (I) (see AAW30501) of transcription factor DPl. Claimed peptides  
 CC (II) (see AAW30504-07) containing one or both of 2 motifs (see  
 CC AAW30502-03) of the DEF box are capable of antagonising the  
 CC heterodimerisation of a DP protein with an E2F protein. Also  
 CC claimed are variants of these peptides, especially containing  
 CC substitutions of residues corresponding to residues 167, 169, 171  
 CC and 175 of DP-1, fusion proteins (III) comprising (I) or (II) and a  
 CC membrane translocation sequence (see AAW30508), expression vectors  
 CC encoding (I)-(III) and host cells. (I)-(III) are used  
 CC therapeutically to induce apoptosis, specifically in tumour or  
 CC cardiovascular cells, either in vivo or in vitro, e.g. for purging  
 CC bone marrow. Surgical stents comprising (I)-(III) are used to  
 CC treat or prevent restenosis in patients who have undergone  
 CC angioplasty. (I)-(III) function by inactivating the DNA-binding  
 CC activity of DP/E2F heterodimers. They are also used as research  
 CC reagents, as positive controls in assays for identifying  
 CC antagonists of DP-1/E2F dimerisation and as immunoassay agents.  
 CC Also described is the use of sequences antisense to nucleic acids  
 CC encoding (I)-(III) to control DP levels in cells, particularly by  
 CC gene therapy. When formulated with cytotoxic or cytostatic agents,  
 CC (I)-(III) enhance cell killing.  
 XX

SQ Sequence 16 AA;

Query Match 100.0%; Score 42; DB 19; Length 16;  
 Best Local Similarity 100.0%; Pred. No. 0.0079;  
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 NVLMAMNII 9  
 DB 7 NVLMAMNII 15

RESULT 3  
 AAW30516  
 ID AAW30516 standard; Peptide; 19 AA.  
 XX  
 AC AAW30516;

26-OCT-1998 (first entry)  
 DT  
 DE DP-1 transcription factor antagonist peptide H2mt2.

XX DP-1; transcription factor; E2F protein; apoptosis;  
 KW cell proliferation; cardiovascular cell; restenosis; tumour;  
 KW surgical stent; therapy.

OS Synthetic.  
 OS Homo sapiens.

XX Key Location/Qualifiers  
 FH Misc-difference 2 /note= "R167A mutation"  
 FT Misc-difference 6 /note= "D171A mutation"

XX WO9828334-A1.

XX 02-JUL-1998.

XX 22-DEC-1997; 97WO-GB03506.

XX 20-DEC-1996; 96GB-0026589.

XX



PA (PROL-) PROLIFIX LTD.  
 XX Bandara LR, La Thangue NB;  
 PI WPI; 1998-377596/32.  
 DR Polypeptide fragments of the DP-1 transcription factor - used for  
 XX inducing apoptosis, specifically in tumour and cardiovascular cells,  
 PT e.g. for preventing restenosis  
 PT e.g. for preventing restenosis  
 XX Example D; Page 26; 55pp; English.  
 PS Peptide H2mt2 is based on peptide H2 (see AAW30504) from the DEF box  
 CC (see AAW30501) of transcription factor DP1. In H2mt2, amino acid  
 CC residues of H2 that correspond to DP1 residues Arg167 and Asp171  
 CC are substituted by Ala residues. H2 is an antagonist of the  
 CC heterodimerisation of DP1 with E2F. H2mt2 retains some, but not  
 CC all, of this antagonistic activity. H2 and other claimed peptides  
 CC (see AAW30504-07) from the DEF box region of DP1 can be used to  
 CC induce apoptosis, specifically in tumour and cardiovascular cells,  
 CC e.g. for the prevention of restenosis.  
 XX Sequence 19 AA;  
 SQ Query Match 100.0%; Score 42; DB 19; Length 19;  
 Best Local Similarity 100.0%; Pred. No. 0.0098;  
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 NVLMAMNII 9  
 DB |||||  
 9 NVLMAMNII 17  
 RESULT 4  
 AAW30504  
 ID AAW30504 standard; Peptide; 19' AA.  
 XX AC AAW30504;  
 XX 26-OCT-1998 (first entry)  
 DT DP-1 transcription factor antagonist peptide H2.  
 DE DP-1; transcription factor; antagonist; E2F protein; apoptosis;  
 XX cell proliferation; cardiovascular cell; restenosis; tumour;  
 KW surgical stent; therapy.  
 KW Synthetic.  
 XX Homo sapiens.  
 OS Key Location/Qualifiers  
 FH Peptide 9..18  
 FT /note= "Claim 3"  
 XX WO9828334-A1.  
 PN 02-JUL-1998.  
 XX 22-DEC-1997; 97WO-GB03506.  
 XX 20-DEC-1996; 96GB-0026589.  
 PR (PROL-) PROLIFIX LTD.  
 XX Bandara LR, La Thangue NB;  
 PI WPI; 1998-377596/32.  
 DR Polypeptide fragments of the DP-1 transcription factor - used for  
 XX inducing apoptosis, specifically in tumour and cardiovascular cells,  
 PT e.g. for preventing restenosis  
 PT e.g. for preventing restenosis  
 XX Claim 4; Page 44; 55pp; English.

XX Peptide H2 comprises amino acid residues 166-184 in the DEF box  
 CC (I) (see AAW30501) of transcription factor DP1. Claimed peptides  
 CC (II) (see AAW30504-07) containing one or both of 2 motifs (see  
 CC AAW30502-03) of the DEF box are capable of antagonising the  
 CC heterodimerisation of a DP protein with an E2F protein. Also  
 CC claimed are variants of these peptides, especially containing  
 CC substitutions of residues corresponding to residues 167, 169, 171  
 CC and 175 of DP-1, fusion proteins (III) comprising (I) or (II) and a  
 CC membrane translocation sequence (see AAW30508), expression vectors  
 CC encoding (I)-(III) and host cells. (I)-(III) are used  
 CC therapeutically to induce apoptosis, specifically in tumour or  
 CC cardiovascular cells, either in vivo or in vitro, e.g. for purging  
 CC bone marrow. Surgical stents comprising (I)-(III) are used to  
 CC treat or prevent restenosis in patients who have undergone  
 CC angioplasty. (I)-(III) function by inactivating the DNA-binding  
 CC activity of DP/E2F heterodimers. They are also used as research  
 CC reagents, as positive controls in assays for identifying  
 CC antagonists of DP-1/E2F dimerisation and as immunoassay agents.  
 CC Also described is the use of sequences antisense to nucleic acids  
 CC encoding (I)-(III) to control DP levels in cells, particularly by  
 CC gene therapy. When formulated with cytotoxic or cytostatic agents,  
 CC (I)-(III) enhance cell killing.  
 XX Sequence 19 AA;  
 SQ Query Match 100.0%; Score 42; DB 19; Length 19;  
 Best Local Similarity 100.0%; Pred. No. 0.0098;  
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 NVLMAMNII 9  
 DB |||||  
 9 NVLMAMNII 17  
 RESULT 5  
 AAW30505  
 ID AAW30505 standard; Peptide; 20 AA.  
 XX AC AAW30505;  
 XX 26-OCT-1998 (first entry)  
 DT DP-1 transcription factor antagonist peptide H3.  
 DE DP-1; transcription factor; antagonist; E2F protein; apoptosis;  
 XX cell proliferation; cardiovascular cell; restenosis; tumour;  
 KW surgical stent; therapy.  
 KW Synthetic.  
 XX Homo sapiens.  
 OS Key Location/Qualifiers  
 FH Peptide 1..10  
 FT /note= "Claim 3"  
 XX WO9828334-A1.  
 PN 02-JUL-1998.  
 XX 22-DEC-1997; 97WO-GB03506.  
 XX 20-DEC-1996; 96GB-0026589.  
 PR (PROL-) PROLIFIX LTD.  
 XX Bandara LR, La Thangue NB;  
 PI WPI; 1998-377596/32.  
 DR Polypeptide fragments of the DP-1 transcription factor - used for  
 XX inducing apoptosis, specifically in tumour and cardiovascular cells,  
 PT e.g. for preventing restenosis  
 PT e.g. for preventing restenosis

Tue Feb 17 11:55:46 2004

PT and arteriosclerosis  
XX Example 3; Page 27; 52pp; Japanese.  
PS  
XX This represents a compound that can inhibit E2F activity. The compound  
CC is of the formula R1 - A - R2 where R1 is an optionally substituted  
CC alkanyl, allyl, hetero-arylcarbonyl, alkoxyarylcarbonyl, aryloxyarylcarbonyl,  
CC hetero-aryloxyarylcarbonyl, or H, R2 is OH, or optionally substituted alkoxy  
CC or amino, and A is an E2F family dimer forming region or DNA binding  
CC region, of at least 12 consecutive amino acids. Compounds of this formula  
CC can be used to inhibit E2F activity, and are useful in the treatment and  
CC prevention of tumours and arteriosclerosis.  
XX Sequence 28 AA;  
SQ Query Match 100.0%; Score 42; DB 19; Length 28;  
Best Local Similarity 100.0%; Pred. No. 0.016; 0; Gaps 0;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 NVLMAMNII 9  
DB 19 NVLMAMNII 27  
RESULT 7  
AAW57055  
ID AAW57055 standard; peptide; 28 AA.  
AC AAW57055;  
XX  
XX 28-AUG-1998 (first entry)  
DT  
XX E2F activity inhibiting compound Ib-3.  
DE  
XX E2F activity; inhibitor; treatment; tumour; arteriosclerosis.  
KW  
XX Synthetic.  
OS  
OS Homo sapiens.  
XX  
XX Key Location/Qualifiers  
FH Modified-site 1 /note= "N-terminal lauroyl"  
FT Modified-site 28  
FT Modified-site /note= "C-terminal amide"  
FT  
XX WO9814474-A1.  
XX  
XX 09-APR-1998.  
PD  
XX 26-SEP-1997; 97WO-JP03442.  
XX  
XX 30-SEP-1996; 96JP-0259432.  
XX  
XX (KYOW ) KYOWA HAKKO KOGYO KK.  
XX  
XX Mizukami T, Shibata K, Yamasaki M, Yoshida T;  
XX  
XX WPI; 1998-240020/21.  
DR  
XX E2F activity inhibitors - for treatment and prevention of tumours  
XX and arteriosclerosis  
XX  
XX Example 7; Page 33; 52pp; Japanese.  
PS  
XX This represents a compound that can inhibit E2F activity. The compound  
CC is of the formula R1 - A - R2 where R1 is an optionally substituted  
CC alkanyl, allyl, hetero-arylcarbonyl, alkoxyarylcarbonyl, aryloxyarylcarbonyl,  
CC hetero-aryloxyarylcarbonyl, or H, R2 is OH, or optionally substituted alkoxy  
CC or amino, and A is an E2F family dimer forming region or DNA binding  
CC region, of at least 12 consecutive amino acids. Compounds of this formula  
CC can be used to inhibit E2F activity, and are useful in the treatment and  
CC prevention of tumours and arteriosclerosis.  
XX

Claim 4; Page 44; 55pp; English.  
XX Peptide H3 comprises amino acid residues 174-193 in the DEF box  
XX (I) (see AAW30501) of transcription factor p1. Claimed peptides  
CC (II) (see AAW30504-07) containing one or both of 2 motifs (see  
CC AAW30502-03) of the DEF box are capable of antagonising the  
CC heterodimerisation of a DP protein with an E2F protein. Also  
CC claimed are variants of these peptides, especially containing  
CC substitutions of residues corresponding to residues 167, 169, 171  
CC and 175 of DP-1, fusion proteins (III) comprising (I) or (II) and a  
CC membrane translocation sequence (see AAW30508), expression vectors  
CC encoding (I)-(III) and host cells. (I)-(III) are used  
CC therapeutically to induce apoptosis, specifically in tumour or  
CC cardiovascular cells, either in vivo or in vitro, e.g. for purging  
CC bone marrow. Surgical stents comprising (I)-(III) are used to  
CC treat or prevent restenosis in patients who have undergone  
CC angioplasty. (I)-(III) function by inactivating the DNA-binding  
CC activity of DP/E2F heterodimers. They are also used as research  
CC reagents, as positive controls in assays for identifying  
CC antagonists of DP-1/E2F dimerisation and as immunoassay agents.  
CC Also described is the use of sequences antisense to nucleic acids  
CC encoding (I)-(III) to control DP levels in cells, particularly by  
CC gene therapy. When formulated with cytotoxic or cytostatic agents,  
CC (I)-(III) enhance cell killing.  
XX Sequence 20 AA;  
SQ Query Match 100.0%; Score 42; DB 19; Length 20;  
Best Local Similarity 100.0%; Pred. No. 0.01; 0; Gaps 0;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 NVLMAMNII 9  
DB 1 NVLMAMNII 9  
RESULT 6  
AAW57051  
ID AAW57051 standard; peptide; 28 AA.  
AC AAW57051;  
XX  
XX 28-AUG-1998 (first entry)  
DT  
XX E2F activity inhibiting compound Ib-1.  
DE  
XX E2F activity; inhibitor; treatment; tumour; arteriosclerosis.  
KW  
XX Synthetic.  
OS  
OS Homo sapiens.  
XX  
XX Key Location/Qualifiers  
FH Modified-site 1 /note= "N-terminal acetyl"  
FT Modified-site 28  
FT Modified-site /note= "C-terminal amide"  
FT  
XX WO9814474-A1.  
XX  
XX 09-APR-1998.  
PD  
XX 26-SEP-1997; 97WO-JP03442.  
XX  
XX 30-SEP-1996; 96JP-0259432.  
XX  
XX (KYOW ) KYOWA HAKKO KOGYO KK.  
XX  
XX Mizukami T, Shibata K, Yamasaki M, Yoshida T;  
XX  
XX WPI; 1998-240020/21.  
DR  
XX E2F activity inhibitors - for treatment and prevention of tumours  
XX

SQ Sequence 28 AA;  
 Query Match 100.0%; Score 42; DB 19; Length 28;  
 Best Local Similarity 100.0%; Pred. No. 0.016;  
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 NVLMANNII 9  
 |||||  
 Db 19 NVLMANNII 27

RESULT 8  
 AAW30507  
 ID AAW30507 standard; Peptide; 30 AA.  
 XX  
 AC AAW30507;  
 DT 26-OCT-1998 (first entry)  
 XX  
 DE DP-1 transcription factor antagonist peptide H7.  
 XX  
 KW DP-1; transcription factor; antagonist; E2F protein; apoptosis;  
 cell proliferation; cardiovascular cell; restenosis; tumour;  
 KW surgical stent; therapy.  
 XX  
 OS Synthetic.  
 OS Homo sapiens.  
 FH Key Location/Qualifiers  
 FT Peptide 3..9  
 FT /note= "Claim 3"  
 FT Peptide 5..15  
 FT /note= "Claim 3"  
 XX  
 PN WO9828334-A1.  
 XX  
 PD 02-JUL-1998.  
 XX  
 PF 22-DEC-1997; 97WO-GB03506.  
 XX  
 PR 20-DEC-1996; 96GB-0026589.  
 XX  
 PA (PROL-) PROLIFIX LTD.  
 XX  
 PI Bandara LR, La Thangue NB;  
 XX  
 DR WPI; 1998-377596/32.  
 XX  
 PT Polypeptide fragments of the DP-1 transcription factor - used for  
 inducing apoptosis, specifically in tumour and cardiovascular cells,  
 e.g. for preventing restenosis  
 XX  
 PS Claim 4; Page 44; 55pp; English.  
 XX

Peptide H7 comprises amino acid residues 170-199 in the DEF box  
 of transcription factor DPL. Claimed peptides  
 (I) (see AAW30501) containing one or both of 2 motifs (see  
 AAW30502-03) of the DEF box are capable of antagonising the  
 heterodimerisation of a DP protein with an E2F protein. Also  
 claimed are variants of these peptides, especially containing  
 substitutions of residues corresponding to residues 167, 169, 171  
 and 175 of DP-1, fusion proteins (III) comprising (I) or (II) and a  
 membrane translocation sequence (see AAW30508), expression vectors  
 encoding (I)-(III) and host cells. (I)-(III) are used  
 therapeutically to induce apoptosis, specifically in tumour or  
 cardiovascular cells, either in vivo or in vitro, e.g. for purging  
 bone marrow. Surgical stents comprising (I)-(III) are used to  
 treat or prevent restenosis in patients who have undergone  
 angioplasty. (I)-(III) function by inactivating the DNA-binding  
 activity of DP/E2F heterodimers. They are also used as research  
 reagents, as positive controls in assays for identifying  
 antagonists of DP-1/E2F dimerisation and as immunoassay agents.  
 Also described is the use of sequences antisense to nucleic acids

CC encoding (I)-(III) to control DP levels in cells, particularly by  
 gene therapy. When formulated with cytotoxic or cytostatic agents,  
 CC (I)-(III) enhance cell killing.  
 XX  
 SQ Sequence 30 AA;  
 Query Match 100.0%; Score 42; DB 19; Length 30;  
 Best Local Similarity 100.0%; Pred. No. 0.017;  
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 NVLMANNII 9  
 |||||  
 Db 5 NVLMANNII 13

RESULT 9  
 AAW30501  
 ID AAW30501 standard; Peptide; 37 AA.  
 XX  
 AC AAW30501;  
 XX  
 DT 26-OCT-1998 (first entry)  
 XX  
 DE DP-1 transcription factor peptide H (DEF box).  
 XX  
 KW DP-1; transcription factor; antagonist; E2F protein; apoptosis;  
 cell proliferation; cardiovascular cell; restenosis; tumour;  
 KW surgical stent; therapy.  
 XX  
 OS Synthetic.  
 OS Homo sapiens.  
 XX  
 PN WO9828334-A1.  
 XX  
 PD 02-JUL-1998.  
 XX  
 PF 22-DEC-1997; 97WO-GB03506.  
 XX  
 PR 20-DEC-1996; 96GB-0026589.  
 XX  
 PA (PROL-) PROLIFIX LTD.  
 XX  
 PI Bandara LR, La Thangue NB;  
 XX  
 DR WPI; 1998-377596/32.  
 XX  
 PT Polypeptide fragments of the DP-1 transcription factor - used for  
 inducing apoptosis, specifically in tumour and cardiovascular cells,  
 e.g. for preventing restenosis  
 XX  
 PS Claim 1; Page 44; 55pp; English.  
 XX

Peptide H (I) comprises residues 163-199, i.e. the DEF box region,  
 of transcription factor DPL. Claimed fragments (II) (see AAW30502-07)  
 of (I) are capable of antagonising the heterodimerisation of a DP  
 protein with an E2F protein. Also claimed are fusion proteins  
 (III) comprising (I) or (II) and a membrane translocation sequence  
 (see AAW30508), expression vectors encoding (I)-(III) and host cells.  
 (I)-(III) are used therapeutically to induce apoptosis,  
 specifically in tumour or cardiovascular cells, either in vivo or in  
 vitro, e.g. for purging bone marrow. Surgical stents comprising  
 (I)-(III) are used to treat or prevent restenosis in patients who  
 have undergone angioplasty. (I)-(III) function by inactivating  
 the DNA-binding activity of DP/E2F heterodimers. They are also  
 used as research reagents, as positive controls in assays for  
 identifying antagonists of DP-1/E2F dimerisation and as immunoassay  
 agents. Also described is the use of sequences antisense to  
 nucleic acids encoding (I)-(III) to control DP levels in cells,  
 particularly by gene therapy. When formulated with cytotoxic  
 or cytostatic agents, (I)-(III) enhance cell killing.  
 XX  
 SQ Sequence 37 AA;

Query Match 100.0%; Score 42; DB 19; Length 37;  
Best Local Similarity 100.0%; Pred. No. 0.022; Mismatches 0; Indels 0; Gaps 0;  
Matches 9; Conservative 0;

QY 1 NVLMANNII 9  
| | | | |  
Db 12 NVLMANNII 20

RESULT 10  
AAW30515  
ID AAW30515 standard; Peptide; 19 AA.

XX AC AAW30515;

XX DT 26-OCT-1998 (first entry)

XX DE DP-1 transcription factor peptide H2mt1.

XX KW DP-1; transcription factor; antagonist; E2F protein; apoptosis;  
KW cell proliferation; cardiovascular cell; restenosis; tumour;  
KW surgical stent; therapy.

XX OS Synthetic.

XX OS Homo sapiens.

XX FH Key Location/Qualifiers

FT Misc-difference 4 /note= "V169A mutation"

FT Misc-difference 10 /note= "V175A mutation"

FT WO9828334-A1.

PN 02-JUL-1998.

PD 22-DEC-1997; 97WO-GB03506.

PF 20-DEC-1996; 96GB-0026589.

PR (PROL-) PROLIFIX LTD.

PA Bandara LR, La Thangue NB;

PI WPI; 1998-377596/32.

XX Polypeptide fragments of the DP-1 transcription factor - used for  
XX inducing apoptosis, specifically in tumour and cardiovascular cells,  
XX e.g. for preventing restenosis

PS Example D; Page 26; 55pp; English.

XX Peptide H2mt1 is based on peptide H2 (see AAW30504) from the DEF box  
XX (see AAW30501) of transcription factor DP1. The H2mt1 peptide, in  
XX which H2 residues corresponding to DP1 residues Val169 and Val175  
XX are substituted by Ala residues, behaves in a similar fashion to  
XX the wild-type H2 peptide in its ability to inactivate E2F site DNA  
XX binding activity in D9 EC cell extracts. H2 is an antagonist of the  
XX heterodimerisation of DP1 with E2F. Thus, the Val-169 and Val-175  
XX residues of H2 play little role in this activity. H2 and other  
XX claimed peptides (see AAW30504-07) from the DEF box region of DP1 can  
XX be used to induce apoptosis, specifically in tumour and  
XX cardiovascular cells, e.g. for the prevention of restenosis.

XX SQ Sequence 19 AA;

Query Match 90.5%; Score 38; DB 19; Length 19;

Best Local Similarity 88.9%; Pred. No. 0.078; Mismatches 1; Indels 0; Gaps 0;  
Matches 8; Conservative 0;

QY 1 NVLMANNII 9

| | | | |

Db 9 NALMANNII 17

RESULT 11  
AAW32163  
ID AAW32163 standard; Protein; 83 AA.  
XX AC AAW32163;  
XX DT 01-FEB-2000 (first entry)

XX DE Soybean DP-1 protein fragment.

XX KW DP-1; soybean; cell cycle regulatory protein;  
KW transcription factor; herbicide.

XX OS Glycine max.

XX FH Key Location/Qualifiers

FT Misc-difference 10 /note= "encoded by GNC"

FT Misc-difference 25 /note= "encoded by ANT"

FT Misc-difference 26 /note= "encoded by GNC"

FT Misc-difference 34 /note= "encoded by GGN"

FT Misc-difference 35 /note= "encoded by CNT"

FT Misc-difference 49 /note= "encoded by NAN"

FT Misc-difference 54 /note= "encoded by ANG"

FT Misc-difference 55 /note= "encoded by NAT"

FT Misc-difference 59 /note= "encoded by NAG"

FT Misc-difference 63 /note= "encoded by NAT"

FT Misc-difference 69 /note= "encoded by GNG"

FT Misc-difference 71 /note= "encoded by NTC"

FT Misc-difference 80 /note= "encoded by CNA"

FT WO953075-A2.

PN 21-OCT-1999.

PD 08-APR-1999; 99WO-US07638.

PF 09-APR-1998; 98US-0081132.

PR (DUPO ) DU PONT DE NEMOURS & CO E I.

XX Klein TM, Morakinyo LO, Odell JT, Sakai H;

XX WPI; 1999-633830/54.

XX N-PSDB; AAZ34579.

XX Plant-derived cell cycle regulatory proteins -

XX Claim 10; Page 41; 44pp; English.

XX This sequence represents 42% of the middle region of soybean cell  
XX cycle regulatory protein DP-1, as deduced from an isolated  
XX cDNA clone (see AAZ34579). The invention relates to nucleic acid  
XX fragments (see AAZ34575-83) encoding plant CDC-16, DP-1, DP-2 and  
XX E2F cell cycle regulatory proteins (see AAY32159-67). It also  
XX relates to the construction of a chimeric gene encoding all or a  
XX portion of the cell cycle regulatory protein, in sense or antisense  
XX orientation, where expression of the chimeric gene results in  
XX production of altered levels of the cell cycle regulatory protein in

CC a transformed host cell. The nucleic acids and proteins may be  
 CC used to facilitate studies of cell cycle regulation in plants,  
 CC provide genetic tools to enhance cell growth in tissue culture,  
 CC increase gene transfer efficiency and provide more stable  
 CC transformations. The proteins may also provide targets to  
 CC facilitate design and/or identification of cell cycle regulatory  
 CC proteins that may be useful as herbicides.

XX SQ Sequence 83 AA;

Query Match 88.1%; Score 37; DB 20; Length 83;  
 Best Local Similarity 88.9%; Pred. No. 0.8;  
 Matches 8; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 NVLMAMNII 9  
 DB 15 NVLMAMNDII 23  
 |||||:|

# RESULT 12

AAW30517  
 ID AAW30517 standard; Peptide; 19 AA.

XX AC AAW30517;

XX DT 26-OCT-1998 (first entry)

XX DE DP-1 transcription factor peptide H2mt3.

XX KW DP-1; transcription factor; antagonist; E2F protein; apoptosis;  
 KW cell proliferation; cardiovascular cell; restenosis; tumour;  
 KW surgical stent; therapy.

XX OS Synthetic.

XX OS Homo sapiens.

XX FH Key Location/Qualifiers  
 FT Misc-difference 8  
 FT /note= "L173R mutation"

FT FT Misc-difference 11  
 FT /note= "L176R mutation"

XX PN WO9828334-A1.

XX PD 02-JUL-1998.

XX PF 22-DEC-1997; 97WO-GB03506.

XX PR 20-DEC-1996; 96GB-0026589.

XX PA (PROL-) PROLIFIX LTD.

XX PI Bandara LR, La Thangue NB;

XX DR WPI; 1998-377596/32.

XX PT Polypeptide fragments of the DP-1 transcription factor - used for  
 PT inducing apoptosis, specifically in tumour and cardiovascular cells,  
 PT e.g. for preventing restenosis

XX PS Example D; Page 26; 55pp; English.

XX CC Peptide H2mt3 is based on peptide H2 (see AAW30504) from the DEF box  
 CC (see AAW30501) of transcription factor DPl. In H2mt3, amino acid  
 CC residues of H2 that correspond to DPl residues Leu173 and Leu176  
 CC are substituted by Arg residues. H2 is an antagonist of the  
 CC heterodimerisation of DPl with E2F. H2mt3 has none of the  
 CC antagonistic activity of H2. H2 and other claimed peptides (see  
 CC AAW30504-07) from the DEF box region of DPl can be used to induce  
 CC apoptosis, specifically in tumour and cardiovascular cells, e.g.  
 CC for the prevention of restenosis.

XX SQ Sequence 19 AA;

Query Match 85.7%; Score 36; DB 19; Length 19;  
 Best Local Similarity 88.9%; Pred. No. 0.22;  
 Matches 8; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 NVLMAMNII 9  
 DB 9 NVLMAMNII 17  
 |||||

# RESULT 13

AAW30511  
 ID AAW30511 standard; Peptide; 14 AA.

XX AC AAW30511;

XX DT 26-OCT-1998 (first entry)

XX DE DP-1 transcription factor peptide H6.

XX KW DP-1; transcription factor; antagonist; E2F protein; apoptosis;  
 KW cell proliferation; cardiovascular cell; restenosis; tumour;  
 KW surgical stent; therapy.

XX OS Synthetic.

XX OS Homo sapiens.

XX PN WO9828334-A1.

XX PD 02-JUL-1998.

XX PF 22-DEC-1997; 97WO-GB03506.

XX PR 20-DEC-1996; 96GB-0026589.

XX PA (PROL-) PROLIFIX LTD.

XX PI Bandara LR, La Thangue NB;

XX DR WPI; 1998-377596/32.

XX PT Polypeptide fragments of the DP-1 transcription factor - used for  
 PT inducing apoptosis, specifically in tumour and cardiovascular cells,  
 PT e.g. for preventing restenosis

XX PS Example C; Page 41; 55pp; English.

XX CC Peptide H6 comprises amino acid residues 167-180 in the DEF box  
 CC region (see AAW30501) of transcription factor DPl. Unlike claimed  
 CC peptides (see AAW30504-07) that contain one or both of 2 motifs (see  
 CC AAW30502-03) of the DPl DEF box, peptide H6 is not capable of  
 CC antagonising the heterodimerisation of a DPl protein with an E2F  
 CC protein. The claimed peptides, their variants and fusion proteins  
 CC can be used to induce apoptosis, specifically in tumour and  
 CC cardiovascular cells, e.g. to prevent restenosis.

XX SQ Sequence 14 AA;

Query Match 81.0%; Score 34; DB 19; Length 14;  
 Best Local Similarity 100.0%; Pred. No. 0.42;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 NVLMAMN 7  
 DB 8 NVLMAMN 14  
 |||||

# RESULT 14

AAW58362  
 ID AAW58362 standard; Protein; 67 AA.

XX AC AAW58362;

XX XX

Tue Feb 17 11:55:46 2004

us-09-900-147-2.rag

```

XX PD 05-NOV-2001 (first entry)
XX DE Human brain expressed single exon probe encoded protein SEQ ID NO: 30467.
XX KW Human; brain expressed exon; gene expression analysis; probe;
XX KW microarray; Alzheimer's disease; multiple sclerosis; schizophrenia;
XX KW epilepsy; cancer.
XX OS Homo sapiens.
XX PN WO200157275-A2.
XX PD 09-AUG-2001.
XX PF 30-JAN-2001; 2001WO-US00667.
XX PR 04-FEB-2000; 2000US-0180312.
XX PR 26-MAY-2000; 2000US-0207456.
XX PR 30-JUN-2000; 2000US-0608408.
XX PR 03-AUG-2000; 2000US-0632366.
XX PR 21-SEP-2000; 2000US-0234687.
XX PR 27-SEP-2000; 2000US-0236359.
XX PR 04-OCT-2000; 2000GB-0024263.
XX FA (MOLE-) MOLECULAR DYNAMICS INC.
XX PI Penn SG, Hanzel DK, Chen W, Rank DR;
XX WPI; 2001-483446/52.
XX DR Single exon nucleic acid probes for analyzing gene expression in human
XX PT brains -
XX PS Example 4; SEQ ID NO: 30467; 650pp + Sequence Listing; English.
XX CC The present invention provides a number of single exon nucleic acid
XX CC probes which are derived from genomic sequences expressed in the human
XX CC brain. They can be used to measure gene expression in brain cell samples,
XX CC which may enable the diagnosis and improved treatment of nervous system,
XX CC diseases such as Alzheimer's disease, multiple sclerosis, schizophrenia,
XX CC epilepsy and cancers. The present sequence is a protein encoded by one of
XX CC the probes of the invention.
XX CC Sequence 67 AA;
XX SQ
XX Query Match 73.8%; Score 31; DB 22; Length 67;
XX Best Local Similarity 55.6%; Pred. No. 14;
XX Matches 5; Conservative 4; Mismatches 0; Indels 0; Gaps 0;
XX QY 1 NVLMAMNII 9
XX DB |:::|:|
XX 43 NIIIIAINII 51
XX RESULT 15
XX AAU72578
XX ID AAU72578 standard; Peptide; 8 AA.
XX AC AAU72578;
XX XX 26-FEB-2002 (first entry)
XX DE Plant dimerisation protein (DP) E2F heterodimerisation domain #1.
XX KW Cell cycle protein; CCP; cell cycle regulation; herbicide;
XX KW plant growth regulator; plant development; abiotic stress; biotic stress;
XX KW nutrient deprivation; pathogen attack; crop yield; motif.
XX XX Glycine max.
XX OS Lycopersicon esculentum.
XX OS Gossypium hirsutum.
XX XX WO200185946-A2.
XX PN
```

```

XX PD 15-NOV-2001.
XX PF 14-MAY-2001; 2001WO-IB01307.
XX PR 12-MAY-2000; 2000US-204045P.
XX PA (CROP-) CROPDESIGN NV.
XX XI Inze D, Boudolf V, De Veylder L, Acosta JAT, Magyar Z;
XX PI WPI; 2002-062249/08.
XX DR New cell cycle protein and nucleic acid molecule encoding it useful for
XX PT regulating cell cycle progression in plants and for identifying
XX PT modulators which are useful as herbicides or plant growth regulators -
XX PS Example 9; Page 93; 316pp; English.
XX CC The invention relates to a novel cell cycle protein (CCP) and the
XX CC polynucleotides encoding them. CCP is useful for identifying a compound
XX CC which modulates the activity of the polypeptide and which binds to the
XX CC polypeptide and an anti-CCP antibody is useful for detecting the presence
XX CC of CCP in a sample. A CCP modulator is useful for modulating the cell
XX CC cycle or growth of a plant such as Arabidopsis thaliana, rice, wheat,
XX CC maize, tomato, alfalfa, oilseed rape, soybean, sunflower and canola.
XX CC CCP nucleic acid and polypeptide molecules are useful as modulating
XX CC agents in regulating cell cycle progression in plants. CCP is useful to
XX CC treat disorders characterised by insufficient or excessive production of
XX CC CCP protein or production of CCP protein forms which have decreased or
XX CC aberrant activity. Compounds that bind to or modulate the activity
XX CC of CCP polypeptide are useful as herbicides or plant growth regulators.
XX CC The polynucleotide is useful for modifying cell fate, plant development,
XX CC plant morphology, biochemistry and/or physiology, the length of the G1,
XX CC S, G2 and/or M phase of the cell cycle of a plant, initiation, promotion,
XX CC stimulation or enhancement of cell division, DNA replication, shoot and root
XX CC seed size, seed development, tuber, fruit, leaf formation, shoot and root
XX CC initiation and/or development, nodule function, dwarfism in plants,
XX CC senescence, tolerance or resistance to stress. CCP, the polynucleotide
XX CC and the anti-CCP antibody are useful in agriculture to modulate the
XX CC protein levels or activity of a protein involved in the cell cycle due
XX CC to environmental conditions, including abiotic stress such as
XX CC cold, nutrient deprivation, heat, drought, salt stress, or biotic
XX CC stress such as pathogen attack, to modulate e.g. enhance crop yields,
XX CC and attenuate plant architecture, plant quality traits, plant
XX CC reproduction and seed development, endoreduplication in storage cells,
XX CC storage tissues and/or storage organs of plants or its parts. CCP is
XX CC useful as an immunogen to generate antibodies. CCP protein is useful to
XX CC screen for naturally occurring CCP substrates. The polynucleotide is
XX CC useful for expressing CCP protein, to detect CCP mRNA, or a genetic
XX CC lesion in a CCP gene and to modulate CCP activity. The present sequence
XX CC represents a motif which may be found in a CCP protein of the invention.
XX SQ
XX Sequence 8 AA;
XX Query Match 69.0%; Score 29; DB 23; Length 8;
XX Best Local Similarity 85.7%; Pred. No. 9.3e+05;
XX Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
XX QY 1 NVLMAMN 7
XX DB |:::|:|
XX 2 NVLMAMD 8
XX RESULT 16
XX AAU77753
XX ID AAU77753 standard; Protein; 88 AA.
XX AC AAU77753;
XX XX 30-OCT-1998 (first entry)
XX DT Pyruvate oxidase protein.
XX DE
```

XX Staphylococcus aureus protein; immune response induction; eye infection;  
 KW antibody production; T-cell immune response; gastrointestinal infection;  
 KW respiratory infection; inhibitor; bacterial infection; cardiac infection;  
 KW central nervous system; kidney infection; urinary tract infection;  
 KW antimicrobial compound identification; broad spectrum antibiotic;  
 XX therapy.

XX Staphylococcus aureus.

XX Key Location/Qualifiers  
 FH Misc-difference 1...88  
 FT /note= "residues designated X are unspecified, and  
 FT represented as Xaa in the specification"

XX EP841394-A2.  
 PN 13-MAY-1998.  
 PD 24-SEP-1997; 97EP-0307485.  
 XX 24-SEP-1996; 96US-0027032.  
 PR (SMIK ) SMITHKLINE BEECHAM CORP.  
 PA (SMIK ) SMITHKLINE BEECHAM PLC.

XX Black MT, Burnham MKR, Hodgson JE, Knowles DJC,  
 PI Lonetto MA, Nicholas RO, Pratt JM, Reichard RW, Rosenberg M;  
 PI Ward JM;

XX WPI; 1998-252940/23.  
 DR N-PSDB; AAV53541.

XX New nucleic acid sequences from Staphylococcus aureus WCHU29 -  
 PT useful in vaccines and for treatment of bacterial infections of e.g.  
 PT respiratory tract and central nervous system

XX Claim 11; Page 372; 390pp; English.

XX This sequence represents a Staphylococcus aureus protein, that based on  
 CC homology with a Lactobacillus plantarum protein, is a pyruvateoxidase  
 CC (EC 1.2.3.3) mutant (lth Pl78S, Sl88N, and A458V mutations) A chain,  
 CC and is encoded by a DNA sequence of the invention.  
 CC The DNA sequences were isolated from Staphylococcus aureus WCHU29  
 CC (NCIB 40771). Host cells containing the DNA sequences are used to  
 CC produce polypeptides or fragments. The proteins are used in the treatment  
 CC of disease, for inducing an immune response by administering them, to  
 CC produce antibody and/or T-cell immune response. Antagonists of the  
 CC proteins are used for the inhibition of bacterial polypeptides.  
 CC Conditions which may be treated include bacterial infections, especially  
 CC respiratory, cardiac, gastrointestinal, central nervous, eye, kidney,  
 CC urinary tract, skin, bones and joints. The proteins can also be used to  
 CC identify antimicrobial compounds which are broad spectrum antibiotics,  
 CC especially useful in the treatment of H. pylori infection.

XX Sequence 88 AA;

Query Match 69.0%; Score 29; DB 19; Length 88;  
 Best Local Similarity 66.7%; Pred. No. 54;  
 Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 NVLMAMNII 9  
 DB 55 NVLVPINII 63  
 |||:||||

RESULT 17  
 ABP38852  
 ID ABP38852 standard; Protein; 89 AA.  
 XX AC ABP38852;  
 XX 24-JUL-2002 (first entry)

XX Staphylococcus epidermidis ORF amino acid sequence SEQ ID NO:3697.  
 DE Staphylococcus epidermidis; open reading frame; ORF; bacterial infection;  
 XX antibacterial; gene therapy.

XX Staphylococcus epidermidis.  
 XX US6380370-B1.  
 XX 30-APR-2002.  
 XX 13-AUG-1998; 98US-0134001.  
 PF 14-AUG-1997; 97US-055779P.  
 PR 08-NOV-1997; 97US-064964P.  
 XX (GENO-) GENOME THERAPEUTICS CORP.

XX Doucette-Stamm LA, Bush D;  
 PI WPI; 2002-381255/41.  
 DR N-PSDB; ABN91397.

XX Novel isolated nucleic acid encoding a Staphylococcus epidermidis  
 PT polypeptide, useful for diagnosing and treating bacterial infections -  
 PS Disclosure; SEQ ID 3697; 267pp; English.

XX ABN90538 to ABN93374 represent Staphylococcus epidermidis open reading  
 CC frame (ORF) nucleic acid sequences which encode the amino acid sequences  
 CC given in ABP35124 to ABP37960. The S. epidermidis sequences have  
 CC antibacterial activity and can be used in gene therapy. The sequences  
 CC can also be used in the diagnosis and treatment of bacterial infections,  
 CC particularly S. epidermidis infections. The sequences can be used to  
 CC screen for compounds able to interfere with the S. epidermidis life  
 CC cycle or inhibit S. epidermidis infection.  
 CC N.B. The sequence data for this patent did not form part of the printed  
 CC specification, but was obtained in electronic format directly from the  
 CC USPTO web site.

XX Sequence 89 AA;  
 SQ Query Match 69.0%; Score 29; DB 23; Length 89;  
 Best Local Similarity 85.7%; Pred. No. 54;  
 Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 2 VLMAMNI 8  
 DB 59 VLMALNI 65  
 ||||:||

RESULT 18  
 ABP38589  
 ID ABP38589 standard; Protein; 100 AA.  
 XX AC ABP38589;  
 XX 24-JUL-2002 (first entry)

XX Staphylococcus epidermidis ORF amino acid sequence SEQ ID NO:3434.  
 DE Staphylococcus epidermidis; open reading frame; ORF; bacterial infection;  
 KW antibacterial; gene therapy.

XX Staphylococcus epidermidis.  
 XX OS US6380370-B1.  
 XX 30-APR-2002.  
 XX 13-AUG-1998; 98US-0134001.

PR 14-AUG-1997; 97US-055779P.  
 PR 08-NOV-1997; 97US-064964P.  
 XX (GENO-) GENOME THERAPEUTICS CORP.  
 PA Doucette-Stamm LA, Bush D;  
 XX WPI; 2002-381255/41.  
 XX N-PSDB; ABN91134.  
 DR Novel isolated nucleic acid encoding a Staphylococcus epidermidis  
 XX polypeptide, useful for diagnosing and treating bacterial infections -  
 PT Disclosure; SEQ ID 3434; 267pp; English.  
 XX AN90538 to ABN93374 represent Staphylococcus epidermidis open reading  
 CC frame (ORF) nucleic acid sequences which encode the amino acid sequences  
 CC given in ABP35124 to ABP37960. The S. epidermidis sequences have  
 CC antibacterial activity and can be used in gene therapy. The sequences  
 CC can also be used in the diagnosis and treatment of bacterial infections,  
 CC particularly S. epidermidis infections. The sequences can be used to  
 CC screen for compounds able to interfere with the S. epidermidis life  
 CC cycle or inhibit S. epidermidis infection.  
 CC N.B. the sequence data for this patent did not form part of the printed  
 CC specification, but was obtained in electronic format directly from the  
 CC USPTO web site.  
 XX SQ Sequence 100 AA;  
 SQ Query Match 69.0%; Score 29; DB 23; Length 100;  
 Best Local Similarity 85.7%; Pred. No. 63;  
 Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;  
 OY 2 VLMAMNI 8  
 DB 49 VLMALNI 55  
 RESULT 19  
 AAW57052  
 ID AAW57052 standard; peptide; 15 AA.  
 XX AC AAW57052;  
 XX DT 28-AUG-1998 (first entry)  
 XX DE E2F activity inhibiting compound Ib-2.  
 XX KW E2F activity; inhibitor; treatment; tumour; arteriosclerosis.  
 XX OS Synthetic.  
 XX OS Homo sapiens.  
 XX FH Key Location/Qualifiers  
 FT Modified-site 1 /note= "N-terminal acetyl"  
 FT Modified-site 15 /note= "C-terminal amide"  
 FT WO9814474-A1.  
 XX PN 09-APR-1998.  
 XX PF 26-SEP-1997; 97WO-JP03442.  
 XX PR 30-SEP-1996; 96JP-0259432.  
 XX PA (KYOW ) KYOWA HAKKO KOGYO KK.  
 XX PI Mizukami T, Shibata K, Yamasaki M, Yoshida T;  
 XX WPI; 1998-240020/21.  
 XX

PT E2F activity inhibitors - for treatment and prevention of tumours  
 PT and arteriosclerosis  
 XX Example 4; Page 28; 52pp; Japanese.  
 XX This represents a compound that can inhibit E2F activity. The compound  
 CC is of the formula R1 - A - R2 where R1 is an optionally substituted  
 CC alkanyl, allyl, hetero-arylcarbonyl, alkoxycarbonyl, aryloxyalkanyl,  
 CC hetero-aryloxyalkanyl, or H, R2 is OH, or optionally substituted alkoxy  
 CC or amino, and A is an E2F family dimer forming region or DNA binding  
 CC region, of at least 12 consecutive amino acids. Compounds of this formula  
 CC can be used to inhibit E2F activity, and are useful in the treatment and  
 CC prevention of tumours and arteriosclerosis.  
 XX SQ Sequence 15 AA;  
 SQ Query Match 66.7%; Score 28; DB 19; Length 15;  
 Best Local Similarity 100.0%; Pred. No. 10;  
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 OY 1 NVLMAM 6  
 DB 10 NVLMAM 15  
 RESULT 20  
 AAB67766  
 ID AAB67766 standard; peptide; 56 AA.  
 XX AC AAB67766;  
 XX DT 11-JUN-2001 (first entry)  
 XX DE Fragment from a wheat E2F-dimerisation partner (DP) protein.  
 XX KW E2F-dimerisation partner; DP protein; E2F transcription factor;  
 KW GI phase; S phase; cell cycle; retinoblastoma protein;  
 KW alter cell proliferation.  
 XX OS Triticum monococcum.  
 XX PN WO200121644-A2.  
 XX PD 29-MAR-2001.  
 XX PF 25-SEP-2000; 2000WO-BF09325.  
 XX PR 24-SEP-1999; 99ES-0002127.  
 XX PR 11-NOV-1999; 99ES-0002474.  
 XX PA (CNSJ ) CONSEJO SUPERIOR INVESTIGACIONES CIENTIF.  
 XX PI Gutierrez-Armenta C, Ramirez-Parra E;  
 XX WPI; 2001-257972/26.  
 XX N-PSDB; AAF80148.  
 XX New isolated, enriched, cell free and/or recombinant nucleic acid  
 PT useful for e.g. altering cell proliferation characteristic such as to  
 PT alter plant cell, organ or tissue size -  
 XX Claim 10; Page 50; 77pp; English.  
 XX AAB67764-68 represent fragments of a wheat E2F-dimerisation partner  
 CC (DP) protein. The protein acts as a plant E2F transcription factor.  
 CC E2F and DP are two proteins that hetero-dimerise to form an active  
 CC transcription factor that regulates G1 to S phase of the cell cycle,  
 CC and later, the expression of genes required for S-phase progression.  
 CC E2F and retinoblastoma protein also interact as a hetero-dimer in  
 CC cells to suppress certain genes. This repression involves binding of  
 CC the retinoblastoma protein to the E2F-DP dimer that is in turn bound  
 CC to sites on DNA through the E2F DNA binding domain. DP proteins can  
 CC be modulated to alter plant cell, organ or tissue shape, and



CC particularly to alter cell proliferation characteristic such as to  
XX alter plant cell, organ or tissue size.

SQ Sequence 56 AA;

Query Match 66.7%; Score 28; DB 22; Length 56;  
Best Local Similarity 55.6%; Pred. No. 52;  
Matches 5; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 1 NVLMAMNII 9  
|:|:|:  
Db 22 NVLIALRVI 30

RESULT 21

AAO00494  
ID AAO00494 standard; Protein; 91 AA.

XX AAO00494;

XX 06-NOV-2001 (first entry)

DT Human polypeptide SEQ ID NO 14386.

XX Human; cytokine; cell proliferation; cell differentiation; gene therapy;  
KW vaccine; peptide therapy; stem cell growth factor; haematopoiesis;  
KW tissue growth factor; immunomodulatory; cancer; leukaemia;  
KW nervous system disorders; arthritis; inflammation.

XX Homo sapiens.

XX WO200164835-A2.

XX 07-SEP-2001.

XX 26-FEB-2001; 2001WO-US04927.

XX 28-FEB-2000; 2000US-0515126.

XX 18-MAY-2000; 2000US-0577409.

XX (HYSE-) HYSEQ INC.

XX Tang YT, Liu C, Drmanac RT;

XX WPI; 2001-514838/56.

XX N-PSDB; AAI80425.

XX Isolated nucleic acids and polypeptides, useful for preventing  
PT diagnosing and treating e.g. leukaemia, inflammation and immune  
PT disorders -

XX Claim 20; SEQ ID NO 14386; 1399pp + Sequence Listing; English.

XX The invention relates to human polynucleotides (AAI79941-AAI93841) and  
CC the encoded proteins (AAO00010-AAO13910) that exhibit activity relating to  
CC cytokine, cell proliferation or cell differentiation or which may induce  
CC production of other cytokines in other cell populations. The  
CC polynucleotides and polypeptides are useful in gene therapy, vaccines or  
CC peptide therapy. The polypeptides have various cytokine-like activities,  
CC e.g. stem cell growth factor activity, haematopoiesis regulating  
CC activity, tissue growth factor activity, immunomodulatory activity and  
CC activin/inhibin activity and may be useful in the diagnosis and/or  
CC treatment of cancer, leukaemia, nervous system disorders, arthritis and  
CC inflammation.

CC Note: The sequence data for this patent did not form part of the printed  
CC specification, but was obtained in electronic format directly from WIPO  
CC at ftp.wipo.int/pub/published\_pct\_sequences.

XX Sequence 91 AA;

Query Match 66.7%; Score 28; DB 22; Length 91;  
Best Local Similarity 55.6%; Pred. No. 94;  
Matches 5; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 1 NVLMAMNII 9  
|:|:|:  
Db 30 NILKSINII 38

RESULT 22  
AAU44499  
ID AAU44499 standard; Protein; 83 AA.

XX AAU44499;

XX 27-FEB-2002 (first entry)

DE Propionibacterium acnes immunogenic protein #5395.

XX SAPHO syndrome; synovitis; acne; pustulosis; hypertosis; osteomyelitis;  
KW uveitis; endophthalmitis; bone; joint; central nervous system; ELISA;  
KW inflammatory lesion; acne vulgaris; enzyme linked immunosorbent assay;  
KW dermatological; osteopathic; neuroprotectant.

XX Propionibacterium acnes.

XX WO200181581-A2.

XX 01-NOV-2001.

XX 20-APR-2001; 2001WO-US12865.

XX 21-APR-2000; 2000US-199047P.

XX 02-JUN-2000; 2000US-208841P.

XX 07-JUL-2000; 2000US-216747P.

XX (CORI-) CORIXA CORP.

XX Skeiky YAW, Persing DH, Mitcham JL, Wang SS, Bhatia A;

XX L'maisonneuve J, Zhang Y, Jen S, Carter D;

XX WPI; 2001-616774/71.

XX N-PSDB; AAS59522.

XX Propionibacterium acnes polypeptides and nucleic acids useful for  
PT vaccinating against and diagnosing infections, especially useful for  
PT treating acne vulgaris -

XX Example 1; SEQ ID No 5694; 1069pp; English.

XX Sequences AAU39105-AAU68017 represent Propionibacterium acnes immunogenic  
CC polypeptides. The proteins and their associated DNA sequences are used in  
CC the treatment, prevention and diagnosis of medical conditions caused by  
CC P. acnes. The disorders include SAPHO syndrome (synovitis, acne,  
CC pustulosis, hypertosis and osteomyelitis), uveitis and endophthalmitis.  
CC P. acnes is also involved in infections of bone, joints and the central  
CC nervous system, however it is particularly involved in the inflammatory  
CC lesions associated with acne vulgaris. A method for detecting the  
CC presence or absence of P. acnes in a patient comprises contacting a  
CC sample with a binding agent that binds to the proteins of the invention  
CC and determining the amount of bound protein in the sample. The  
CC polypeptides may be used as antigens in the production of antibodies  
CC specific for P. acnes proteins. These antibodies can be used to  
CC downregulate expression and activity of P. acnes polypeptides and  
CC therefore treat P. acnes infections. The antibodies may also be used as  
CC diagnostic agents for determining P. acnes presence, for example, by  
CC enzyme linked immunosorbent assay (ELISA).

CC Note: The sequence data for this patent did not form part of the printed  
CC specification, but was obtained in electronic format directly from WIPO  
CC at ftp.wipo.int/pub/published\_pct\_sequences.

XX Sequence 83 AA;

Query Match 64.3%; Score 27; DB 22; Length 83;  
Best Local Similarity 62.5%; Pred. No. 1.4e+02;  
Matches 5; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Tue Feb 17 11:55:46 2004

us-09-900-147-2-2.rag

QY 1 NVLMAMNI 8  
Db 9 NVLVNWNV 16  
RESULT 23  
AAU21311  
ID AAU21311 standard; Protein; 48 AA.  
XX AAU21311;  
AC  
DT 18-DEC-2001 (first entry)  
XX Human novel foetal antigen, SEQ ID NO 1555.  
XX Human; foetal tissue antigen; antiinflammatory; neuroprotective;  
KW immunomodulator; cardiovascular; cytostatic; nephrothropic;  
KW cardiovascular; autoimmune disease; rheumatoid arthritis;  
KW hyperproliferative disorder; breast neoplasm; cancer;  
KW cardiovascular disorder; cardiac arrest; cerebrovascular disorder;  
KW cerebral ischaemia; angiogenesis; nervous system disorder;  
KW Alzheimer's disease; infection; ocular disorder; corneal infection;  
KW wound healing; epithelial cell proliferation; food additive.  
XX  
OS Homo sapiens.  
XX  
XX WO200155312-A2.  
XX 02-AUG-2001.  
XX  
XX 17-JAN-2001; 2001WO-US01321.  
XX  
XX 31-JAN-2000; 2000US-0179065.  
PR 04-FEB-2000; 2000US-0180628.  
PR 24-FEB-2000; 2000US-0184664.  
PR 02-MAR-2000; 2000US-0186350.  
PR 16-MAR-2000; 2000US-0189874.  
PR 17-MAR-2000; 2000US-0190076.  
PR 18-APR-2000; 2000US-0198123.  
PR 19-MAY-2000; 2000US-0205515.  
PR 07-JUN-2000; 2000US-0209467.  
PR 28-JUN-2000; 2000US-0214886.  
PR 30-JUN-2000; 2000US-0215135.  
PR 07-JUL-2000; 2000US-0216647.  
PR 07-JUL-2000; 2000US-0216880.  
PR 11-JUL-2000; 2000US-0217487.  
PR 11-JUL-2000; 2000US-0217496.  
PR 14-JUL-2000; 2000US-0218290.  
PR 26-JUL-2000; 2000US-0220963.  
PR 26-JUL-2000; 2000US-0220964.  
PR 14-AUG-2000; 2000US-0224518.  
PR 14-AUG-2000; 2000US-0224519.  
PR 14-AUG-2000; 2000US-0225213.  
PR 14-AUG-2000; 2000US-0225214.  
PR 14-AUG-2000; 2000US-0225266.  
PR 14-AUG-2000; 2000US-0225267.  
PR 14-AUG-2000; 2000US-0225268.  
PR 14-AUG-2000; 2000US-0225270.  
PR 14-AUG-2000; 2000US-0225447.  
PR 14-AUG-2000; 2000US-0225757.  
PR 14-AUG-2000; 2000US-0225758.  
PR 14-AUG-2000; 2000US-0225759.  
PR 18-AUG-2000; 2000US-0226279.  
PR 22-AUG-2000; 2000US-0226681.  
PR 22-AUG-2000; 2000US-0226868.  
PR 22-AUG-2000; 2000US-0227182.  
PR 23-AUG-2000; 2000US-0227009.  
PR 30-AUG-2000; 2000US-0228924.  
PR 01-SEP-2000; 2000US-0229287.  
PR 01-SEP-2000; 2000US-0229343.  
PR 01-SEP-2000; 2000US-0229344.  
PR 01-SEP-2000; 2000US-0229345.

PR 05-SEP-2000; 2000US-0229509.  
PR 05-SEP-2000; 2000US-0229513.  
PR 06-SEP-2000; 2000US-0230437.  
PR 06-SEP-2000; 2000US-0230438.  
PR 08-SEP-2000; 2000US-0231242.  
PR 08-SEP-2000; 2000US-0231243.  
PR 08-SEP-2000; 2000US-0231244.  
PR 08-SEP-2000; 2000US-0231413.  
PR 08-SEP-2000; 2000US-0231414.  
PR 08-SEP-2000; 2000US-0232080.  
PR 08-SEP-2000; 2000US-0232081.  
PR 12-SEP-2000; 2000US-0231968.  
PR 14-SEP-2000; 2000US-0232397.  
PR 14-SEP-2000; 2000US-0232398.  
PR 14-SEP-2000; 2000US-0232399.  
PR 14-SEP-2000; 2000US-0232400.  
PR 14-SEP-2000; 2000US-0232401.  
PR 14-SEP-2000; 2000US-0233063.  
PR 14-SEP-2000; 2000US-0233064.  
PR 14-SEP-2000; 2000US-0233065.  
PR 21-SEP-2000; 2000US-0234223.  
PR 21-SEP-2000; 2000US-0234274.  
PR 25-SEP-2000; 2000US-0234997.  
PR 25-SEP-2000; 2000US-0234998.  
PR 26-SEP-2000; 2000US-0235484.  
PR 27-SEP-2000; 2000US-0235834.  
PR 27-SEP-2000; 2000US-0235836.  
PR 29-SEP-2000; 2000US-0236327.  
PR 29-SEP-2000; 2000US-0236367.  
PR 29-SEP-2000; 2000US-0236368.  
PR 29-SEP-2000; 2000US-0236369.  
PR 29-SEP-2000; 2000US-0236370.  
PR 02-OCT-2000; 2000US-0236802.  
PR 02-OCT-2000; 2000US-0237037.  
PR 02-OCT-2000; 2000US-0237038.  
PR 02-OCT-2000; 2000US-0237039.  
PR 02-OCT-2000; 2000US-0237040.  
PR 13-OCT-2000; 2000US-0239935.  
PR 13-OCT-2000; 2000US-0239937.  
PR 20-OCT-2000; 2000US-0240960.  
PR 20-OCT-2000; 2000US-0241221.  
PR 20-OCT-2000; 2000US-0241785.  
PR 20-OCT-2000; 2000US-0241786.  
PR 20-OCT-2000; 2000US-0241787.  
PR 20-OCT-2000; 2000US-0241808.  
PR 20-OCT-2000; 2000US-0241809.  
PR 20-OCT-2000; 2000US-0241826.  
PR 01-NOV-2000; 2000US-0244617.  
PR 08-NOV-2000; 2000US-0246474.  
PR 08-NOV-2000; 2000US-0246475.  
PR 08-NOV-2000; 2000US-0246476.  
PR 08-NOV-2000; 2000US-0246477.  
PR 08-NOV-2000; 2000US-0246478.  
PR 08-NOV-2000; 2000US-0246523.  
PR 08-NOV-2000; 2000US-0246524.  
PR 08-NOV-2000; 2000US-0246525.  
PR 08-NOV-2000; 2000US-0246526.  
PR 08-NOV-2000; 2000US-0246527.  
PR 08-NOV-2000; 2000US-0246528.  
PR 08-NOV-2000; 2000US-0246532.  
PR 08-NOV-2000; 2000US-0246533.  
PR 08-NOV-2000; 2000US-0246609.  
PR 08-NOV-2000; 2000US-0246610.  
PR 08-NOV-2000; 2000US-0246611.  
PR 08-NOV-2000; 2000US-0246613.  
PR 17-NOV-2000; 2000US-0249207.  
PR 17-NOV-2000; 2000US-0249208.  
PR 17-NOV-2000; 2000US-0249209.  
PR 17-NOV-2000; 2000US-0249210.  
PR 17-NOV-2000; 2000US-0249211.  
PR 17-NOV-2000; 2000US-0249212.  
PR 17-NOV-2000; 2000US-0249213.  
PR 17-NOV-2000; 2000US-0249214.  
PR 17-NOV-2000; 2000US-0249215.

```

PR 17-NOV-2000; 2000US-0249216.
PR 17-NOV-2000; 2000US-0249217.
PR 17-NOV-2000; 2000US-0249218.
PR 17-NOV-2000; 2000US-0249244.
PR 17-NOV-2000; 2000US-0249245.
PR 17-NOV-2000; 2000US-0249264.
PR 17-NOV-2000; 2000US-0249265.
PR 17-NOV-2000; 2000US-0249297.
PR 17-NOV-2000; 2000US-0249299.
PR 17-NOV-2000; 2000US-0249300.
PR 01-DEC-2000; 2000US-0250160.
PR 01-DEC-2000; 2000US-0250391.
PR 05-DEC-2000; 2000US-0251030.
PR 05-DEC-2000; 2000US-0251988.
PR 05-DEC-2000; 2000US-0256719.
PR 06-DEC-2000; 2000US-0251479.
PR 08-DEC-2000; 2000US-0251856.
PR 08-DEC-2000; 2000US-0251868.
PR 08-DEC-2000; 2000US-0251869.
PR 08-DEC-2000; 2000US-0251989.
PR 08-DEC-2000; 2000US-0251990.
PR 11-DEC-2000; 2000US-0254097.
PR 05-JAN-2001; 2001US-0259678.
XX (HUMA-) HUMAN GENOME SCI INC.
PA
XX
XX Rosen CA, Barash SC, Ruben SM;
PI
XX
XX WPI; 2001-488782/53.
XX N-PSDB; AAS34131.
XX
XX New polynucleotides and polypeptides for diagnosing, treating,
PT preventing or prognosing e.g. diseases or disorders of the nervous,
PT musculoskeletal, excretory, gastrointestinal, reproductive, and
PT respiratory systems
XX
XX Claim 11; SEQ ID No 1555; 642pp; English.
XX
XX The invention relates to novel nucleic acids encoding novel human foetal
CC antigens. The nucleic acids and proteins are used to prevent, treat (e.g.
CC by gene therapy) or ameliorate a medical condition in e.g. humans, mice,
CC rabbits, goats, horses, cats, dogs, chickens or sheep. They
CC are also used in diagnosing a pathological condition or susceptibility
CC to a pathological condition. The antibodies to the antigens can also
CC be used in alleviating symptoms associated with the disorders and in
CC diagnostic immunoassays e.g. radioimmunoassays or enzyme linked
CC immunosorbent assays (ELISA). Disorders which are diagnosed or treated
CC include autoimmune diseases e.g. rheumatoid arthritis,
CC hyperproliferative disorders e.g. neoplasms of the breast or liver,
CC cardiovascular disorders e.g. cardiac arrest, cerebrovascular disorders
CC e.g. cerebral ischaemia, angiogenesis, nervous system disorders e.g.
CC Alzheimer's disease, infections caused by bacteria, viruses and fungi
CC and ocular disorders e.g. corneal infection. The polypeptides can also
CC be used to aid wound healing and epithelial cell proliferation, to
CC prevent skin aging due to sunburn, to maintain organs before
CC transplantation, for supporting cell culture of primary tissues, to
CC regenerate tissues and in chemotaxis. The polypeptides can also be used
CC as a food additive or preservative to increase or decrease storage
CC capabilities, fat content, lipid, protein, carbohydrate, vitamins,
CC minerals, cofactors and other nutritional components. Numerous
CC examples of diseases and disorders treated by the nucleic acids and
CC proteins are given in the specification. The present sequence
Query Match 61.9%; Score 26; DB 22; Length 48;
Best Local Similarity 62.5%; Pred. No. 1.2e+02;
Matches 5; Conservative 2; Mismatches 1; Indels 0; Gaps 0;
QY 2 VLMAMNII 9
Db 17 VLAAMNLL 24
RESULT 24

```

---

```

AAG75292
ID AAG75292 standard; Protein; 55 AA.
XX
XX AAG75292;
AC
XX
XX 03-SEP-2001 (first entry)
DT
XX
XX Human colon cancer antigen protein SEQ ID NO:6056.
DE
XX
XX Human; colon cancer; colon cancer antigen; diagnosis; detection;
XX colorectal carcinoma; chromosome 1.
XX
XX Homo sapiens.
OS
XX
XX WO200122920-A2.
PN
XX
XX 05-APR-2001.
PD
XX
XX 28-SEP-2000; 2000WO-US26524.
PF
XX
XX 29-SEP-1999; 99US-0157137.
PR
XX
XX 03-NOV-1999; 99US-0163280.
PR
XX
XX (HUMA-) HUMAN GENOME SCI INC.
PA
XX
XX Ruben SM, Barash SC, Birse CE, Rosen CA;
PI
XX
XX WPI; 2001-235357/24.
XX N-PSDB; AAH34697.
XX
XX Nucleic acids encoding 4277 human colon cancer-associated polypeptides,
PT useful for preventing, diagnosing and/or treating colorectal cancers -
PT
XX
XX Claim 11; Page 7535-7536; 9803pp; English.
XX
XX AAH32943 to AAH37195 and AAG73514 to AAG77788 represent human colon
CC cancer-associated nucleic acid molecules (N) and proteins (P), where
CC the proteins are collectively known as colon cancer antigens. The colon
CC cancer antigens have cytostatic activity and can be used in gene
CC therapy and vaccine production. N and P may be used in the prevention,
CC diagnosis and treatment of diseases associated with inappropriate P
CC expression. For example, N and P may be used to treat disorders
CC associated with decreased expression by rectifying mutations or deletions
CC in a patient's genome that affect the activity of P by expressing
CC inactive proteins or to supplement the patients own production of P.
CC Additionally, N may be used to produce the colon cancer-associated Ps
CC by inserting the nucleic acids into a host cell and culturing the cell
CC to express the proteins. N and P can be used in the prevention, diagnosis
CC and treatment of colorectal carcinomas and cancers. AAH37196 to AAH37204
CC and AAB77789 represent sequences used in the exemplification of the
CC present invention.
CC N.B. Pages 666 to 682 and page 7053 of the sequence listing were
CC missing at time of publication, meaning no sequences are present for
CC SEQ ID NO:1027 to 1052, 7921 and 7922.
XX
XX Sequence 55 AA;
SQ
Query Match 61.9%; Score 26; DB 22; Length 55;
Best Local Similarity 62.5%; Pred. No. 1.4e+02;
Matches 5; Conservative 2; Mismatches 1; Indels 0; Gaps 0;
QY 2 VLMAMNII 9
Db 45 VLLAKNIV 52
RESULT 25
ABP03723
ID ABP03723 standard; Protein; 58 AA.
XX
XX ABP03723;
AC
XX
XX 25-JUN-2002 (first entry)
DT

```

XX DE Human ORFX protein sequence SEQ ID NO:7428.

XX Human; open reading frame; ORFX; gene therapy; cancer; cirrhosis;

KW hyperproliferative disorder; psoriasis; benign tumour; haemorrhage;

KW degenerative disorder; osteoarthritis; neurodegenerative disorder;

KW cardiovascular disease; diabetes mellitus; systemic lupus erythematosus;

KW hypertension; hypothyroidism; cholesterol ester storage disease;

KW immune deficiency; immune disorder; infectious disease;

KW autoimmune disorder; rheumatoid arthritis; autoimmune thyroiditis;

KW myasthenia gravis.

XX Homo sapiens.

XX WO200192523-A2.

XX 06-DEC-2001.

XX 29-MAY-2001; 2001WO-US10836.

XX 30-MAY-2000; 2000US-206132P.

XX 29-AUG-2000; 2000US-228716P.

XX (CURA-) CURAGEN CORP.

XX Shimkets RA, Leach MD;

XX WPI; 2002-106308/14.

XX N-PSDB; ABN19475.

XX Novel human polypeptides and polynucleotides useful for diagnosing,

XX preventing and treating cardiovascular disease, neurodegenerative,

XX hyperproliferative disorders and autoimmune disorders

XX Disclosure; SEQ ID 7428; 1037pp; English.

XX The present invention describes substantially purified human proteins

CC (referred to as open reading frame, ORFX, where X is 1-11491 (see Table 1

CC in the specification). ABN15762 to ABN27252 encode the human ORFX

CC proteins given in ABP0010 to ABP11500. ORFX proteins are useful for

CC treating or preventing a pathology associated with an ORFX-associated

CC disorder in humans, and in the manufacture of a medicament for treating a

CC syndrome associated with ORFX-associated disorder. ORFX polynucleotide

CC sequences can be used in gene therapy. ORFX sequences can be used in the

CC treatment of cancer, hyperproliferative disorders, cirrhosis of liver,

CC psoriasis, benign tumours, keloid, degenerative disorders, haemorrhage,

CC osteoarthritis, neurodegenerative disorders, disorders related to organ

CC transplantation, cardiovascular diseases, diabetes mellitus, systemic

CC lupus erythematosus, hypertension, hypothyroidism, cholesterol ester

CC storage disease, various immune deficiencies and disorders, infectious

CC diseases, autoimmune disorders such as multiple sclerosis, rheumatoid

CC arthritis, autoimmune thyroiditis, myasthenia gravis, graft-versus-host

CC disease and autoimmune inflammatory eye disease. ORFX proteins are also

CC useful for treating burns, incisions, ulcers, for treating osteoporosis,

CC bone degenerative disorders, or periodontal disease, and for gut

CC protection or regeneration and treatment of lung or liver fibrosis,

CC reperfusion injury in various tissues and conditions resulting from

CC systemic cytokine damage.

CC N.B. The sequence data for this patent did not form part of the printed

CC specification, but was obtained in electronic format directly from WIPO

CC at ftp.wipo.int/pub/published\_pct\_sequences.

XX SQ Sequence 58 AA;

Query Match 61.9%; Score 26; DB 23; Length 58;

Best Local Similarity 55.6%; Pred. NO. 1.5e+02;

Matches 5; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 1 NVLMANNII 9

8 NIQKSMNII 16

Db

RESULT 26

ABBI7490

ID ABB17490 standard; Protein; 62 AA.

XX AC ABB17490;

XX 23-JAN-2002 (first entry)

XX Human nervous system related polypeptide SEQ ID NO 6147.

XX Human; nootropic; neuroprotective; cytostatic; dermatological; virucide;

KW immunosuppressive; antiinflammatory; anti-HIV; antibacterial; vulnery;

KW antiparkinsonian; antiskilling; antianaemic; antiarthritic; cancer;

KW antirheumatic; hepatoprotective; cerebroprotective; antinflammatory;

KW antiallergic; antidiabetic; antitumor; anticonvulsant; antifungal;

KW antiparasitic; cardiant; immune disorder; cardiovascular disorder;

KW neurological disease; infection; nephrotropic; gene therapy; vaccine.

XX Homo sapiens.

XX WO200159063-A2.

XX 16-AUG-2001.

XX 17-JAN-2001; 2001WO-US01334.

XX 31-JAN-2000; 2000US-0179065.

XX 04-FEB-2000; 2000US-0180628.

XX 24-FEB-2000; 2000US-0184664.

XX 02-MAR-2000; 2000US-0186350.

XX 16-MAR-2000; 2000US-0189874.

XX 17-MAR-2000; 2000US-0190076.

XX 18-APR-2000; 2000US-0198123.

XX 19-MAY-2000; 2000US-0205515.

XX 07-JUN-2000; 2000US-0209467.

XX 28-JUN-2000; 2000US-0214886.

XX 30-JUN-2000; 2000US-0215135.

XX 07-JUL-2000; 2000US-0216647.

XX 07-JUL-2000; 2000US-0216880.

XX 11-JUL-2000; 2000US-0217487.

XX 11-JUL-2000; 2000US-0217496.

XX 14-JUL-2000; 2000US-0218290.

XX 26-JUL-2000; 2000US-0220963.

XX 26-JUL-2000; 2000US-0220964.

XX 14-AUG-2000; 2000US-0224518.

XX 14-AUG-2000; 2000US-0224519.

XX 14-AUG-2000; 2000US-0225213.

XX 14-AUG-2000; 2000US-0225214.

XX 14-AUG-2000; 2000US-0225266.

XX 14-AUG-2000; 2000US-0225267.

XX 14-AUG-2000; 2000US-0225268.

XX 14-AUG-2000; 2000US-0225270.

XX 14-AUG-2000; 2000US-0225447.

XX 14-AUG-2000; 2000US-0225757.

XX 14-AUG-2000; 2000US-0225758.

XX 14-AUG-2000; 2000US-0225759.

XX 18-AUG-2000; 2000US-0226279.

XX 22-AUG-2000; 2000US-0226681.

XX 22-AUG-2000; 2000US-0226868.

XX 22-AUG-2000; 2000US-0227182.

XX 23-AUG-2000; 2000US-0227009.

XX 30-AUG-2000; 2000US-0228924.

XX 01-SEP-2000; 2000US-0229287.

XX 01-SEP-2000; 2000US-0229343.

XX 01-SEP-2000; 2000US-0229344.

XX 01-SEP-2000; 2000US-0229345.

XX 05-SEP-2000; 2000US-0229509.

XX 05-SEP-2000; 2000US-0229513.

XX 06-SEP-2000; 2000US-0230437.

XX 06-SEP-2000; 2000US-0230438.

XX 08-SEP-2000; 2000US-0231242.

XX 08-SEP-2000; 2000US-0231243.

XX 08-SEP-2000; 2000US-0231244.

PR 08-SEP-2000; 2000US-02311413.  
PR 08-SEP-2000; 2000US-02314114.  
PR 08-SEP-2000; 2000US-02320810.  
PR 08-SEP-2000; 2000US-02320811.  
PR 12-SEP-2000; 2000US-02319668.  
PR 14-SEP-2000; 2000US-02323397.  
PR 14-SEP-2000; 2000US-02323398.  
PR 14-SEP-2000; 2000US-02323399.  
PR 14-SEP-2000; 2000US-02324001.  
PR 14-SEP-2000; 2000US-02330633.  
PR 14-SEP-2000; 2000US-02330634.  
PR 14-SEP-2000; 2000US-02330655.  
PR 21-SEP-2000; 2000US-02342223.  
PR 21-SEP-2000; 2000US-02342274.  
PR 25-SEP-2000; 2000US-02349997.  
PR 25-SEP-2000; 2000US-02349998.  
PR 26-SEP-2000; 2000US-02354884.  
PR 27-SEP-2000; 2000US-02358334.  
PR 27-SEP-2000; 2000US-02358336.  
PR 29-SEP-2000; 2000US-02363277.  
PR 29-SEP-2000; 2000US-02363668.  
PR 29-SEP-2000; 2000US-02363669.  
PR 29-SEP-2000; 2000US-02363700.  
PR 02-OCT-2000; 2000US-0236802.  
PR 02-OCT-2000; 2000US-0237037.  
PR 02-OCT-2000; 2000US-0237038.  
PR 02-OCT-2000; 2000US-0237039.  
PR 02-OCT-2000; 2000US-0237040.  
PR 13-OCT-2000; 2000US-02399335.  
PR 13-OCT-2000; 2000US-02399337.  
PR 20-OCT-2000; 2000US-02409600.  
PR 20-OCT-2000; 2000US-0241785.  
PR 20-OCT-2000; 2000US-0241786.  
PR 20-OCT-2000; 2000US-0241787.  
PR 20-OCT-2000; 2000US-0241808.  
PR 20-OCT-2000; 2000US-0241809.  
PR 20-OCT-2000; 2000US-0241826.  
PR 20-OCT-2000; 2000US-0242221.  
PR 01-NOV-2000; 2000US-0244617.  
PR 08-NOV-2000; 2000US-0246474.  
PR 08-NOV-2000; 2000US-0246475.  
PR 08-NOV-2000; 2000US-0246476.  
PR 08-NOV-2000; 2000US-0246477.  
PR 08-NOV-2000; 2000US-0246478.  
PR 08-NOV-2000; 2000US-0246523.  
PR 08-NOV-2000; 2000US-0246524.  
PR 08-NOV-2000; 2000US-0246525.  
PR 08-NOV-2000; 2000US-0246526.  
PR 08-NOV-2000; 2000US-0246527.  
PR 08-NOV-2000; 2000US-0246528.  
PR 08-NOV-2000; 2000US-0246532.  
PR 08-NOV-2000; 2000US-0246609.  
PR 08-NOV-2000; 2000US-0246610.  
PR 08-NOV-2000; 2000US-0246611.  
PR 08-NOV-2000; 2000US-0246613.  
PR 17-NOV-2000; 2000US-0249207.  
PR 17-NOV-2000; 2000US-0249208.  
PR 17-NOV-2000; 2000US-0249209.  
PR 17-NOV-2000; 2000US-0249210.  
PR 17-NOV-2000; 2000US-0249211.  
PR 17-NOV-2000; 2000US-0249212.  
PR 17-NOV-2000; 2000US-0249213.  
PR 17-NOV-2000; 2000US-0249214.  
PR 17-NOV-2000; 2000US-0249215.  
PR 17-NOV-2000; 2000US-0249217.  
PR 17-NOV-2000; 2000US-0249218.  
PR 17-NOV-2000; 2000US-0249219.  
PR 17-NOV-2000; 2000US-0249244.  
PR 17-NOV-2000; 2000US-0249245.  
PR 17-NOV-2000; 2000US-0249264.  
PR 17-NOV-2000; 2000US-0249265.

PR 17-NOV-2000; 2000US-0249297.  
PR 17-NOV-2000; 2000US-0249299.  
PR 17-NOV-2000; 2000US-0249300.  
PR 01-DEC-2000; 2000US-0250391.  
PR 01-DEC-2000; 2000US-0250391.  
PR 01-DEC-2000; 2000US-0251160.  
PR 05-DEC-2000; 2000US-0251030.  
PR 05-DEC-2000; 2000US-0251988.  
PR 05-DEC-2000; 2000US-0256719.  
PR 06-DEC-2000; 2000US-0251479.  
PR 08-DEC-2000; 2000US-0251856.  
PR 08-DEC-2000; 2000US-0251868.  
PR 08-DEC-2000; 2000US-0251869.  
PR 08-DEC-2000; 2000US-0251989.  
PR 08-DEC-2000; 2000US-0251990.  
PR 11-DEC-2000; 2000US-0254097.  
PR 05-JAN-2001; 2001US-0259678.  
XX  
XX (HUMA-) HUMAN GENOME SCI INC.  
XX  
XX Rosen CA, Barash SC, Ruben SM;  
XX  
XX WPI; 2001-541565/60.  
XX N-PSDB; ABA13816.  
XX  
XX Nucleic acids encoding 3224 human nervous system antigen polypeptides,  
XX useful for preventing, diagnosing and/or treating nervous system  
XX cancers and metastases -  
XX  
XX Claim 11; SEQ ID NO 6147; 1701pp + Sequence Listing; English.  
XX  
XX The invention relates to novel genes (ABAI1004-ABAI21534) and proteins  
XX (ABAI4678-ABAI8001) useful for preventing, treating or ameliorating  
XX medical conditions e.g. by protein or gene therapy. The genes are  
XX isolated from a range of human tissues disclosed in the specification.  
XX The nucleic acids, proteins, antibodies and (ant)agonists are useful  
XX in the diagnosis, treatment and prevention of: (a) cancer, e.g. breast  
XX and ovarian cancer and other cancers of the adrenal gland, bone, bone  
XX marrow, breast, gastrointestinal tract, liver, lung, or urogenital;  
XX (b) immune disorders e.g. Addison's disease, allergies, autoimmune  
XX haemolytic anaemia, autoimmune thyroiditis, diabetes mellitus, Crohn's  
XX disease, multiple sclerosis, rheumatoid arthritis and ulcerative  
XX colitis; (c) cardiovascular disorders such as myocardial ischaemias;  
XX (d) wound healing; (e) neurological diseases e.g. cerebral anoxia and  
XX epilepsy; and (f) infectious diseases such as viral, bacterial, fungal  
XX and parasitic infections.  
XX Note: The sequence data for this patent did not form part of the  
XX printed specification, but was obtained in electronic format directly  
XX from WIPO at ftp.wipo.int/pub/published\_pct\_sequences.  
XX  
SQ Sequence 62 AA;  
Query Match 61.9%; Score 26; DB 22; Length 62;  
Best Local Similarity 50.0%; Pred. No. 1.7e+02;  
Matches 4; Conservative 3; Mismatches 1; Indels 0; Gaps 0;  
Qy 1 NVLMAMNI 8  
Db 44 NILLILNI 51  
RESULT 27  
AAO11790  
ID AAO11790 standard; Protein; 71 AA.  
XX  
XX AAO11790;  
XX  
XX 06-NOV-2001 (first entry)  
XX  
XX Human polypeptide SEQ ID NO 25682.  
XX  
XX Human; cytokine; cell proliferation; cell differentiation; gene therapy;  
XX vaccine; peptide therapy; stem cell growth factor; haematopoiesis;  
XX tissue growth factor; immunomodulatory; cancer; leukaemia;

XX nervous system disorders; arthritis; inflammation.  
 XX Homo sapiens.  
 XX WO200164835-A2.  
 XX 07-SEP-2001.  
 XX 26-FEB-2001; 2001WO-US04927.  
 XX 28-FEB-2000; 2000US-0515126.  
 XX 18-MAY-2000; 2000US-0577409.  
 XX (HYSE-) HYSEQ INC.  
 XX Tang YT, Liu C, Drmanac RT;  
 XX WPI; 2001-514838/56.  
 XX N-PSDB; AAI91721.  
 XX Isolated nucleic acids and polypeptides, useful for preventing  
 XX diagnosing and treating e.g. leukaemia, inflammation and immune  
 XX disorders -  
 XX Claim 20; SEQ ID NO 25682; 1399pp + Sequence Listing; English.  
 XX The invention relates to human polynucleotides (AAI79941-AAI93841) and  
 XX the encoded proteins (AAO00010-AAO13910) that exhibit activity elating to  
 XX cytokine, cell proliferation or cell differentiation or which may induce  
 XX production of other cytokines in other cell populations. The  
 XX polynucleotides and polypeptides are useful in gene therapy, vaccines or  
 XX peptide therapy. The polypeptides have various cytokine-like activities,  
 XX e.g. stem cell growth factor activity, haematopoiesis regulating  
 XX activity, tissue growth factor activity, immunomodulatory activity and  
 XX activin/inhibin activity and may be useful in the diagnosis and/or  
 XX treatment of cancer, leukaemia, nervous system disorders, arthritis and  
 XX inflammation.  
 XX Note: The sequence data for this patent did not form part of the printed  
 XX specification, but was obtained in electronic format directly from WIPO  
 XX at ftp.wipo.int/pub/published\_pct\_sequences.  
 XX Sequence 71 AA;  
 XX Query Match 61.9%; Score 26; DB 22; Length 71;  
 XX Best Local Similarity 44.4%; Pred. No. 2e+02;  
 XX Matches 4; Conservative 3; Mismatches 2; Indels 0; Gaps 0;  
 XX QY 1 NVLMAMNII 9  
 XX Db 63 NLLQLMNLV 71  
 XX  
 XX RESULT 28  
 XX ABP03425  
 XX ID ABP03425 standard; Protein; 72 AA.  
 XX AC ABP03425;  
 XX XX 24-JUN-2002 (first entry)  
 XX DE Human ORFX protein sequence SEQ ID NO:6832.  
 XX Human; open reading frame; ORFX; gene therapy; cancer; cirrhosis;  
 XX hyperproliferative disorder; psoriasis; benign tumour; haemorrhage;  
 XX degenerative disorder; osteoarthritis; neurodegenerative disorder;  
 XX cardiovascular disease; diabetes mellitus; systemic lupus erythematosus;  
 XX hypertension; hypothyroidism; cholesterol ester storage disease;  
 XX immune deficiency; immune disorder; infectious disease;  
 XX autoimmune disorder; rheumatoid arthritis; autoimmune thyroiditis;  
 XX myasthenia gravis.  
 XX Homo sapiens.  
 XX

PN WO200192523-A2.  
 XX 06-DEC-2001.  
 XX 29-MAY-2001; 2001WO-US10836.  
 XX 30-MAY-2000; 2000US-206132P.  
 XX 29-AUG-2000; 2000US-228716P.  
 XX (CURA-) CURAGEN CORP.  
 XX Shimkets RA, Leach MD;  
 XX WPI; 2002-106308/14.  
 XX N-PSDB; AAI9177.  
 XX Novel human polypeptides and polynucleotides useful for diagnosing,  
 XX preventing and treating cardiovascular disease, neurodegenerative,  
 XX hyperproliferative disorders and autoimmune disorders -  
 XX Disclosure; SEQ ID 6832; 1037pp; English.  
 XX The present invention describes substantially purified human proteins  
 XX (referred to as open reading frame, ORFX, where X is 1-11491 (see Table 1  
 XX in the specification). AAI91762 to AAI92752 encode the human ORFX  
 XX proteins given in ABP00010 to ABP11500. ORFX proteins are useful for  
 XX treating or preventing a pathology associated with an ORFX-associated  
 XX syndrome in humans, and in the manufacture of a medicament for treating a  
 XX disorder associated with ORFX-associated disorder. ORFX polynucleotide  
 XX sequences can be used in gene therapy. ORFX sequences can be used in the  
 XX treatment of cancer, hyperproliferative disorders, cirrhosis of liver,  
 XX psoriasis, benign tumours, keloid, degenerative disorders, haemorrhage,  
 XX osteoarthritis, neurodegenerative disorders, diabetes mellitus, systemic  
 XX transplantation, cardiovascular diseases, hypothyroidism, infectious  
 XX storage disease, various immune deficiencies and disorders, rheumatoid  
 XX diseases, autoimmune disorders such as multiple sclerosis, rheumatoid  
 XX arthritis, autoimmune thyroiditis, myasthenia gravis, graft-versus-host  
 XX disease and autoimmune inflammatory eye disease. ORFX proteins are also  
 XX useful for treating burns, incisions, ulcers, for treating osteoporosis,  
 XX bone degenerative disorders, or periodontal disease, and for gut  
 XX protection or regeneration and treatment of lung or liver fibrosis,  
 XX reperfusion injury in various tissues and conditions resulting from  
 XX systemic cytokine damage.  
 XX N.B. The sequence data for this patent did not form part of the printed  
 XX specification, but was obtained in electronic format directly from WIPO  
 XX at ftp.wipo.int/pub/published\_pct\_sequences.  
 XX Sequence 72 AA;  
 XX Query Match 61.9%; Score 26; DB 23; Length 72;  
 XX Best Local Similarity 71.4%; Pred. No. 2e+02;  
 XX Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;  
 XX QY 3 LMAMNII 9  
 XX Db 24 LMAVNLI 30  
 XX  
 XX RESULT 29  
 XX AAW20232  
 XX ID AAW20232 standard; Protein; 79 AA.  
 XX AC AAW20232;  
 XX XX 09-JUL-1997 (first entry)  
 XX DE H. pylori cytoplasmic protein, 22542803.aa.  
 XX Cytoplasmic; vaccine; prevention; treatment; infection; envelope;  
 XX identification; binding compound; bacterium; life cycle; activator;  
 XX bacteria; inhibitor; duodenal ulcer disease; chronic gastritis;  
 XX diagnosis.

```

XX OS Helicobacter pylori.
XX PN WO9640893-A1.
XX PD 19-DEC-1996.
XX PF 06-JUN-1996; 96WO-US09122.
XX PR 01-APR-1996; 96US-0630405.
XX PR 07-JUN-1995; 95US-0487032.
XX PA (ASTR ) ASTRA AB.
XX PI Berglindh OT, Smith D, Mellgaerd BL;
XX DR WPI; 1997-052306/05.
XX DR N-PSDB; AAT67454.
XX PT Helicobacter pylori nucleic acid sequences and related
XX PT polypeptide(s) - useful for vaccines to treat or prevent H. pylori
XX PT infection, and to detect Helicobacter
XX PS Claim 61; Page 440-41; 1481pp; English.
XX CC This sequence represents a H. pylori cytoplasmic protein.
XX CC The protein may be used in a vaccine to prevent or treat H. pylori
XX CC infection or to identify H. pylori polypeptide binding compounds,
XX CC useful as potential H. pylori life cycle activators or inhibitors.
XX CC The genomic sequence of H. pylori (ATCC 55679) was determined from
XX CC overlapping contigs generated by mechanically shearing the bacterial
XX CC DNA. The sequences were analysed for ORF of at least 180 nucleotides,
XX CC and the predicted coding regions defined by computer evaluation. To
XX CC identify likely H. pylori antigens for vaccine development, the amino
XX CC acid sequences predicted from various ORF were analysed for significant
XX CC homology to other known or exported membrane proteins. Having identified
XX CC and determined the sequences of interest, particular regions can be
XX CC isolated from H. pylori by PCR amplification for recombinant polypeptide
XX CC production, e.g. in E. coli hosts.
XX SQ Sequence 79 AA;
Query Match 61.9%; Score 26; DB 18; Length 79;
Best Local Similarity 62.5%; Pred. No. 2.2e+02;
Matches 5; Conservative 2; Mismatches 1; Indels 0; Gaps 0;
QY 1 NVLMAMNI 8
Db ||||| ::
15 NVLMATDV 22
RESULT 30
AAAY10972
XX ID AAY10972 standard; Protein; 96 AA.
XX AC AAY10972;
XX DT 08-JUN-1999 (first entry)
XX DE H. pylori ORF 03eel1215_22542803_fl_7 secreted protein.
XX KW Vaccine; probe; diagnostic; ORF; cell envelope protein;
XX KW secreted protein; cellular protein.
XX OS Helicobacter pylori.
XX PN WO9818323-A1.
XX PD 07-MAY-1998.
XX PF 28-OCT-1997; 97WO-US19575.
XX PR 14-JUL-1997; 97US-0891928.

```

```

PR 28-OCT-1996; 96US-0739150.
PR 06-DEC-1996; 96US-0759739.
XX PA (ASTR ) ASTRA AB.
XX PI Alm RA, Smith D;
XX DR WPI; 1998-271811/24.
XX DR N-PSDB; AAX30439.
XX PT Helicobacter pylori nucleic acids and proteins - used to develop
XX PT products for the detection, prevention and treatment of H. pylori
XX PT infections
XX PS Claims 27, 31; Page 176; 279pp; English.
XX CC Recombinant or substantially pure preparations of H. pylori polypeptides
XX CC are disclosed, together with the nucleic acids encoding them. In all,
XX CC 73 ORFs are shown. The proteins are variously cell envelope proteins,
XX CC secreted proteins or other cellular proteins. Vaccines containing the
XX CC nucleic acids or proteins are claimed, as are probes containing at least
XX CC 8 nucleotides from the nucleic acid sequences. The vaccines are useful
XX CC for treating or reducing the risk of H. pylori infections, and the
XX CC probes can be used diagnostically for detecting the presence of
XX CC Helicobacter in a sample. The products are also of use in screening
XX CC for compounds having the ability to interfere with the H. pylori life
XX CC cycle or to inhibit H. pylori infection.
XX SQ Sequence 96 AA;
Query Match 61.9%; Score 26; DB 19; Length 96;
Best Local Similarity 62.5%; Pred. No. 2.8e+02;
Matches 5; Conservative 2; Mismatches 1; Indels 0; Gaps 0;
QY 1 NVLMAMNI 8
Db ||||| ::
15 NVLMATDV 22
RESULT 31
AAM00188
XX ID AAM00188 standard; Peptide; 14 AA.
XX AC AAM00188;
XX DT 01-OCT-2001 (first entry)
XX DE Human protein fragment SEQ ID NO: 728.
XX KW Human; single nucleotide polymorphism; SNP; paternity test;
XX KW forensic test; aberrant protein expression.
XX OS Homo sapiens.
XX PN WO200151670-A2.
XX PD 19-JUL-2001.
XX PF 05-JAN-2001; 2001WO-US00322.
XX PR 07-JAN-2000; 2000US-0174962.
XX PA (CURA-) CURAGEN CORP.
XX PI Shimkets RA, Leach MD;
XX DR WPI; 2001-451871/48.
XX DR N-PSDB; AAH89297.
XX PT Isolated human polynucleotides containing single nucleotide
XX PT polymorphisms, useful for the treatment and diagnosis of e.g. cancer,
XX PT infection and diabetes -

```

PS Disclosure; Page 314; 475pp; English.

XX The present invention relates to human nucleic acids containing single  
CC nucleotide polymorphisms (SNPs). These can be used in forensic and  
CC paternity tests, and to aid in the treatment of diseases associated with  
CC aberrant protein expression, including cancer, amyloidosis, diabetes,  
CC Alzheimer's disease, Down's syndrome, oedema, lupus (SLE), vasculitis,  
CC glomerulonephritis, haemolytic anaemia, thrombocytopaenia, arthritis,  
CC meningitis, muscular disorders, dementia, hypercalcaemia, tuberculous  
CC sclerosis, male infertility, hypercalcaemia, blood pressure disorders,  
CC osteoporosis, pathogenic infections, hypercholesterolaemia, obesity or  
CC autoimmunity. The present sequence is a peptide encoded by a  
CC polymorphism-containing oligonucleotide fragment of the invention.

XX Sequence 14 AA;

Query Match 59.5%; Score 25; DB 22; Length 14;  
Best Local Similarity 62.5%; Pred. No. 45;  
Matches 5; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Qy 1 NVLMAMNI 8  
Db 7 NYLQALNI 14

RESULT 32  
ID AAM38530 standard; Protein; 25 AA.

XX AAM38530;

XX 17-OCT-2001 (first entry)

XX Peptide #12567 encoded by probe for measuring placental gene expression.

XX Probe; microarray; human; placenta; antenatal diagnosis;  
XX genetic disorder.

XX Homo sapiens.

XX WO200157272-A2.

XX 09-AUG-2001.

XX 30-JAN-2001; 2001WO-US00663.

XX 04-FEB-2000; 2000US-0180312.

XX 26-MAY-2000; 2000US-0207456.

XX 30-JUN-2000; 2000US-0608408.

XX 03-AUG-2000; 2000US-0632366.

XX 21-SEP-2000; 2000US-0234687.

XX 27-SEP-2000; 2000US-0236359.

XX 04-OCT-2000; 2000GB-0024263.

XX (MOLE-) MOLECULAR DYNAMICS INC.

XX Penn SG, Hanzel DK, Chen W, Rank DR;

XX WPI; 2001-488897/53.

XX Human genome-derived single exon nucleic acid probes useful for  
XX analyzing gene expression in human placenta -

XX Claim 27; SEQ ID No 38799; 654pp; English.

XX The present invention relates to single exon nucleic acid probes (SNP;  
CC see AAI31315-AAI57546). The present sequence is a peptide encoded by one  
CC such probe. The probes are useful for producing a microarray for  
CC predicting, measuring and displaying gene expression in samples derived  
CC from human placenta. The probes are useful for antenatal diagnosis of  
CC human genetic disorders.

XX Sequence 25 AA;

Query Match 59.5%; Score 25; DB 22; Length 25;  
Best Local Similarity 33.3%; Pred. No. 91;  
Matches 3; Conservative 5; Mismatches 1; Indels 0; Gaps 0;

Qy 1 NVLMAMNI 9  
Db 15 NILMSLQLL 23

RESULT 33  
ID AAR63496 standard; Peptide; 32 AA.

XX AAR63496;

XX 25-MAR-2003 (updated)

DT 25-MAY-1995 (first entry)

XX Rat REC20 serotonin receptor extracellular loop 1 antigenic peptide.

XX Serotonin; receptor; transmembrane; domain; kinase; phosphorylation;  
KW sensory; motor; behaviour; central nervous system; CNS; superfamily;  
KW G-protein; ligand-gated; ion channel; subfamily; human; rat; amplify;  
KW primer; PCR; amplification; brain; hypothalamus; indolamine; drug;  
KW hypothalamus; therapeutic; neurological; pathology; dementia; insomnia;  
KW Parkinson's disease; eating disorder; anxiety; migraine; headache.

XX Synthetic.

XX Key Location/Qualifiers

FT Misc-difference 8 /note= "undetermined amino acid"

XX WO9421670-A1.

XX 29-SEP-1994.

XX 15-MAR-1994; 94WO-US02839.

XX 15-MAR-1993; 93US-0031538.

XX (SCRI ) SCRIPPS RES INST.

XX Erlander MG, Lovenberg TW, Sutcliffe JG;

XX WPI; 1994-316932/39.

XX New serotonin receptors and corresp. DNA and antibodies - useful  
XX in diagnosis and treatment of neurological processes and  
XX pathologies.

XX Claim 45; Page 174; 198pp; English.

XX A series of synthesised peptides (AAR63492-9) based on the amino acid  
XX sequence of the novel rat serotonin receptor REC20. The peptides were  
XX used either in the generation of antibodies for the diagnosis of the  
XX receptor or in the detection of materials binding to the receptors.  
XX Serotonin receptors belong to at least two protein superfamilies:  
XX G-protein-associated receptors containing 7 TMDs (including  
XX 5-HT1a/b/c/d/e, 5-HT2) and ligand-gated ion channel receptors with 4  
XX TMDs (5-HT3). The serotonin receptors presented in the patent represent  
XX prototypes that fall into three new serotonin subfamily classifications:  
XX 5-HT1e-like (rat MR77 - AAQ72271 and human MR77 - AAQ72272), 5-HT5  
XX (subdivided into 5-HT5alpha, rat REC17 - AAQ72269 and 5-HT5beta, rat  
XX MR22 - AAQ72270) and 5-HT6 (rat REC20 - AAQ72273). The genes for the  
XX receptors were cloned by amplification based on conserved amino acid  
XX sequences found in the TMDs. Degenerate primers were constructed to  
XX these sequences and used in two rounds of nested PCR amplification on a  
XX rat brain hypothalamic cDNA template. In the second round of  
XX amplification, degenerate primers corresponding to conserved residues  
XX found only in indolamine-binding receptors were used to amplify only  
XX those receptors. The resultant fragments were cloned into pBluescript



CC vectors. The plasmid fragments were used to screen a rat hypothalamus  
 CC cDNA library to obtain full length clones. The nucleic acids molecules  
 CC and their corresponding proteins may be used in methods for determining  
 CC ligand binding activity, detecting and altering expression of serotonin  
 CC receptors, drug screening and therapeutic treatments involving human  
 CC serotonin receptors. Serotonin receptors are also known to participate  
 CC in neurological processes. Pathologies including dementia, Parkinson's  
 CC disease, eating disorders, pathological anxiety, migraine, headaches,  
 CC insomnia and other conditions.  
 CC (Updated on 25-MAR-2003 to correct PN field.)  
 XX Sequence 32 AA;  
 SQ Query Match 59.5%; Score 25; DB 15; Length 32;  
 Best Local Similarity 44.4%; Pred. No. 1.2e+02;  
 Matches 4; Conservative 1; Mismatches 1; Indels 0; Gaps 0;  
 QY 1 NVLMANNII 9  
 Db 22 NVFIAMDVM 30  
 |||:::  
 |||:::  
 RESULT 34  
 AAY76229  
 ID AAY76229 standard; Protein; 35 AA.  
 XX  
 AC AAY76229;  
 XX  
 DT 23-MAR-2000 (first entry)  
 XX  
 DE Fragment of human secreted protein encoded by gene 2.  
 XX  
 KW Human; secreted protein; cancer; tumour; developmental abnormality;  
 KW foetal deficiency; blood disorder; immune system disorder; inflammation;  
 KW autoimmune disease; allergy; Alzheimer's disease; cognitive disorder;  
 KW schizophrenia; arthritis; aschma; psoriasis; sepsis; skin disorder;  
 KW atherosclerosis; diabetes; cardiovascular disorder; kidney disorder;  
 KW digestive disorder; endocrine disorder; infection; AIDS; leukaemia;  
 KW therapy.  
 XX  
 OS Homo sapiens.  
 XX  
 XX WO9958660-A1.  
 XX  
 PD 18-NOV-1999.  
 XX  
 XX 06-MAY-1999; 99WO-US09847.  
 XX  
 PR 12-MAY-1998; 98US-0085093.  
 PR 12-MAY-1998; 98US-0085094.  
 PR 12-MAY-1998; 98US-0085105.  
 PR 12-MAY-1998; 98US-0085180.  
 PR 18-MAY-1998; 98US-0085906.  
 PR 18-MAY-1998; 98US-0085920.  
 PR 18-MAY-1998; 98US-0085921.  
 PR 18-MAY-1998; 98US-0085922.  
 PR 18-MAY-1998; 98US-0085923.  
 PR 18-MAY-1998; 98US-0085924.  
 PR 18-MAY-1998; 98US-0085928.  
 PR 18-MAY-1998; 98US-0085925.  
 PR 18-MAY-1998; 98US-0085927.  
 XX  
 PA (HUMA-) HUMAN GENOME SCI INC.  
 XX  
 PI Ruben SM, Florence K, Ni J, Rosen CA, Carter KC, Moore PA;  
 PI Olsen HS, Shi Y, Young PE, Wei F, Brewer LA, Soppet DR;  
 PI Lafleur DW, Endress GA, Ebner R;  
 XX  
 DR WPI; 2000-062296/05.  
 XX  
 XX New isolated human genes and the secreted polypeptides they encode,  
 PT useful for diagnosis and treatment of e.g. cancers, neurological  
 PT disorders, immune diseases, inflammation or blood disorders -

XX Disclosure; Page 417; 475pp; English.  
 XX  
 CC AAZ65250 to AAZ65350 represent 97 isolated human secreted protein genes.  
 CC AAY76124 to AAY76223 are the secreted proteins encoded by the 97 human  
 CC genes. This sequence represents a fragment of one of the human secreted  
 CC proteins. The genes and their corresponding secreted polypeptides are  
 CC useful for preventing, treating or ameliorating medical conditions,  
 CC e.g. by protein or gene therapy. Also pathological conditions can be  
 CC diagnosed by determining the amount of the new polypeptides in a sample  
 CC or by determining the presence of mutations in the new genes. Specific  
 CC uses are described for each of the 97 genes, based on which tissues they  
 CC are most highly expressed in, and include developing products for the  
 CC diagnosis or treatment of cancer, tumours, developmental abnormalities  
 CC and foetal deficiencies, blood disorders, diseases of the immune system,  
 CC autoimmune diseases, inflammation, allergies, Alzheimer's and cognitive  
 CC disorders, schizophrenia, arthritis, asthma, psoriasis, sepsis, skin  
 CC disorders, atherosclerosis, diabetes, cardiovascular disorders, kidney  
 CC disorders, digestive/endocrine disorders, infections and AIDS. The  
 CC polypeptides are also useful for identifying their binding partners.  
 CC The sequences shown in AAY76224 to AAY76424 represent fragments of the  
 CC secreted proteins.  
 XX  
 SQ Sequence 35 AA;  
 Query Match 59.5%; Score 25; DB 21; Length 35;  
 Best Local Similarity 83.3%; Pred. No. 1.4e+02;  
 Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;  
 QY 4 MANNII 9  
 Db 11 MALNII 16  
 |||:::  
 |||:::  
 RESULT 35  
 ABG47854  
 ID ABG47854 standard; Peptide; 51 AA.  
 XX  
 AC ABG47854;  
 XX  
 DT 25-FEB-2003 (first entry)  
 XX  
 DE Human liver peptide, SEQ ID No 26502.  
 XX  
 KW Human; liver; cirrhosis; hyperlipoproteinaemia; hyperlipidaemia;  
 KW hypercholesterolaemia; coronary heart disease.  
 XX  
 OS Homo sapiens.  
 XX  
 XX WO200157273-A2.  
 XX  
 PD 09-AUG-2001.  
 XX  
 PF 30-JAN-2001; 2001WO-US00664.  
 XX  
 PR 04-FEB-2000; 2000US-0180312.  
 PR 26-MAY-2000; 2000US-0207456.  
 PR 30-JUN-2000; 2000US-0608408.  
 PR 03-AUG-2000; 2000US-0632366.  
 PR 21-SEP-2000; 2000US-0234687.  
 PR 27-SEP-2000; 2000US-0236359.  
 PR 04-OCT-2000; 2000GB-0024263.  
 XX  
 PA (MOLE-) MOLECULAR DYNAMICS INC.  
 XX  
 PI Penn SG, Hanzel DK, Chen W, Rank DR;  
 PI  
 XX WPI; 2001-488898/53.  
 DR  
 XX Human genome-derived single exon nucleic acid probes useful for  
 PT analysing gene expression in human adult liver -  
 XX  
 PS Claim 27; SEQ ID No 26502; 658pp; English.

XX The invention relates to a single exon nucleic acid probe (SENP) (I) for  
CC measuring human gene expression in a sample derived from human adult  
CC liver, comprising one of 13109 defined nucleotide sequences given in the  
CC specification (or complements/ fragments). The probe hybridises at high  
CC stringency to a nucleic acid molecule expressed in the human adult  
CC liver. (I) may be used for predicting, measuring and displaying gene  
CC expression in samples derived from human adult liver. The genes  
CC identified may be involved in genetic liver diseases such as cirrhosis,  
CC hyperlipoproteinaemia, hyperlipidaemia and hypercholesterolaemia which  
CC is associated with coronary heart disease. ABC47348-ABG59930 represent  
CC human liver single exon encoded peptides of the invention.  
CC Note: The sequence information for this patent does not appear in the  
CC printed specification but was obtained in electronic format directly  
CC from WIPO at ftp.wipo.int/pub/published\_pct\_sequences.  
XX  
SQ Sequence 51 AA;  
  
Query Match 59.5%; Score 25; DB 22; Length 51;  
Best Local Similarity 71.4%; Pred. No. 2.2e+02;  
Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;  
  
QY 3 LMANMII 9  
Db 17 LMANMNTL 23  
||||| :  
17 LMANMNTL 23  
  
RESULT 36  
ABB27831  
ID ABB27831 standard; Peptide; 51 AA.  
XX  
AC ABB27831;  
XX  
DT 01-FEB-2002 (first entry)  
XX  
DE Human peptide #482 encoded by breast cell single exon nucleic acid probe.  
XX  
KW Human; microarray; single exon probe; gene expression; breast;  
KW disease; cancer.  
XX  
OS Homo sapiens.  
XX  
PN WO200157271-A2.  
XX  
PD 09-AUG-2001.  
XX  
PF 30-JAN-2001; 2001WO-US00662.  
XX  
PR 04-FEB-2000; 2000US-0180312.  
XX  
PR 26-MAY-2000; 2000US-0207456.  
XX  
PR 30-JUN-2000; 2000US-0608408.  
XX  
PR 03-AUG-2000; 2000US-0632366.  
XX  
PR 21-SEP-2000; 2000US-0234687.  
XX  
PR 27-SEP-2000; 2000US-0236359.  
XX  
PR 04-OCT-2000; 2000GB-0024263.  
XX  
PA (MOLE-) MOLECULAR DYNAMICS INC.  
XX  
PI Penn SG, Hanzel DK, Chen W, Rank DR;  
XX  
PT WPI; 2001-496933/54.  
XX  
DR New spatially-addressable set of single exon nucleic acid probes,  
XX useful for measuring gene expression in sample derived from human  
XX breast, comprises number of single exon nucleic acid probes -  
XX  
PS Claim 27; SEQ ID NO 10799; 327pp + sequence listing; English.  
XX  
CC The invention relates to a spatially-addressable set of single exon  
CC nucleic acid probes for measuring gene expression in a sample derived  
CC from human breast and Et 474 cells. The method involves contacting  
CC the probes with a collection of detectably labelled nucleic acids  
CC derived from mRNA of human breast, and then measuring the label

CC bound to each probe of the microarray. The probes are useful for  
CC verifying the expression of regions of genomic DNA predicted to  
CC encode proteins. They are useful for gene discovery, and for  
CC determining predisposition and/or prognosing breast disease. Gene  
CC expression analysis is useful for assessing the toxicity of chemical  
CC agents on cells. The microarray of this invention presents a far greater  
CC diversity of probes for measuring gene expression, with far less bias  
CC than expressed sequence tag microarrays. The method is suitable for  
CC rapid production of functional information from genomic sequence. The  
CC present sequence is a peptide encoded by a single exon nucleic acid  
CC probe of the invention.  
CC Note: The sequence data for this patent did not form part of the  
CC printed specification, but was obtained in electronic format directly  
CC from WIPO at ftp.wipo.int/pub/published\_pct\_sequences.  
XX  
SQ Sequence 51 AA;  
  
Query Match 59.5%; Score 25; DB 22; Length 51;  
Best Local Similarity 71.4%; Pred. No. 2.2e+02;  
Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;  
  
QY 3 LMANMII 9  
Db 17 LMANMNTL 23  
||||| :  
17 LMANMNTL 23  
  
RESULT 37  
ABB33002  
ID ABB33002 standard; Peptide; 51 AA.  
XX  
AC ABB33002;  
XX  
DT 04-FEB-2002 (first entry)  
XX  
DE Peptide #508 encoded by human foetal liver single exon probe.  
XX  
KW Human; foetal liver; gene expression; single exon nucleic acid probe.  
XX  
OS Homo sapiens.  
XX  
PN WO200157277-A2.  
XX  
PD 09-AUG-2001.  
XX  
PF 30-JAN-2001; 2001WO-US00669.  
XX  
PR 04-FEB-2000; 2000US-0180312.  
XX  
PR 26-MAY-2000; 2000US-0207456.  
XX  
PR 30-JUN-2000; 2000US-0608408.  
XX  
PR 03-AUG-2000; 2000US-0632366.  
XX  
PR 21-SEP-2000; 2000US-0234687.  
XX  
PR 27-SEP-2000; 2000US-0236359.  
XX  
PR 04-OCT-2000; 2000GB-0024263.  
XX  
PA (MOLE-) MOLECULAR DYNAMICS INC.  
XX  
PI Penn SG, Hanzel DK, Chen W, Rank DR;  
XX  
PT WPI; 2001-483447/52.  
XX  
DR Human genome-derived single exon nucleic acid probes useful for  
XX analyzing gene expression in human foetal liver -  
XX  
PS Claim 27; SEQ ID NO 25637; 639pp + sequence listing; English.  
XX  
CC The invention relates to a single exon nucleic acid probe for  
CC measuring human gene expression in a sample derived from human foetal  
CC liver. The single exon nucleic acid probes may be used for predicting,  
CC measuring and displaying gene expression in samples derived from human  
CC foetal liver. The present sequence is a peptide encoded by a single exon  
CC nucleic acid probe of the invention.  
CC Note: The sequence data for this patent did not form part of the  
CC printed specification, but was obtained in electronic format directly

CC from WIPO at ftp.wipo.int/pub/published\_pct\_sequences.

XX Sequence 51 AA;

SQ Query Match 59.5%; Score 25; DB 22; Length 51;  
Best Local Similarity 71.4%; Pred. No. 2.2e+02;  
Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 3 LMAMNII 9

Db 17 LMAMNTL 23

RESULT 38

ABB18475

ID ABB18475 standard; Protein; 51 AA.

XX AC ABB18475;

XX 23-JAN-2002 (first entry)

XX DT

XX DE Protein #474 encoded by probe for measuring heart cell gene expression.

XX KW Human; gene expression; heart; microarray; vascular system;

XX KW Cardiovascular disease; hypertension; cardiac arrhythmia;

XX KW congenital heart disease.

XX OS Homo sapiens.

XX PN WO200157274-A2.

XX PD 09-AUG-2001.

XX PF 30-JAN-2001; 2001WO-US00666.

XX PR 04-FEB-2000; 2000US-0180312.

XX PR 26-MAY-2000; 2000US-0207456.

XX PR 30-JUN-2000; 2000US-0608408.

XX PR 03-AUG-2000; 2000US-0632366.

XX PR 21-SEP-2000; 2000US-0234687.

XX PR 27-SEP-2000; 2000US-0236359.

XX PR 04-OCT-2000; 2000GB-0024263.

XX PA (MOLE-) MOLECULAR DYNAMICS INC.

XX PI Penn SG, Hanzel DK, Chen W, Rank DR;

XX WPI; 2001-488899/53.

XX PT Single exon nucleic acid probes for analyzing gene expression in human

XX PT hearts -

XX PS Claim 15; SEQ ID NO 20245; 530pp; English.

XX CC The present invention relates to single exon nucleic acid probes for

CC measuring human gene expression in a sample derived from human heart (see

CC ABA21535-ABA41305). The present sequence is a protein encoded by one such

CC probe. The probes may be used for predicting, measuring and displaying

CC gene expression in samples derived from the human heart via microarrays.

CC By measuring gene expression, the probes are useful for predicting,

CC diagnosing, grading, staging, monitoring and prognosing diseases of the

CC human heart and vascular system e.g. cardiovascular disease,

CC hypertension, cardiac arrhythmias and congenital heart disease.

CC Note: The sequence data for this patent did not form part of the printed

CC specification, but was obtained in electronic format directly from WIPO

CC at ftp.wipo.int/pub/published\_pct\_sequences.

XX SQ Sequence 51 AA;

XX Query Match 59.5%; Score 25; DB 22; Length 51;

XX Best Local Similarity 71.4%; Pred. No. 2.2e+02;

XX Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 3 LMAMNII 9

Db 17 LMAMNTL 23

RESULT 39

AAM53800

ID AAM53800 standard; Protein; 51 AA.

XX AC AAM53800;

XX 05-NOV-2001 (first entry)

XX DT

XX DE Human brain expressed single exon probe encoded protein SEQ ID NO: 25905.

XX KW Human; brain expressed exon; gene expression analysis; probe;

XX KW microarray; Alzheimer's disease; multiple sclerosis; schizophrenia;

XX KW epilepsy; cancer.

XX OS Homo sapiens.

XX PN WO200157275-A2.

XX PD 09-AUG-2001.

XX PF 30-JAN-2001; 2001WO-US00667.

XX PR 04-FEB-2000; 2000US-0180312.

XX PR 26-MAY-2000; 2000US-0207456.

XX PR 30-JUN-2000; 2000US-0608408.

XX PR 03-AUG-2000; 2000US-0632366.

XX PR 21-SEP-2000; 2000US-0234687.

XX PR 27-SEP-2000; 2000US-0236359.

XX PR 04-OCT-2000; 2000GB-0024263.

XX PA (MOLE-) MOLECULAR DYNAMICS INC.

XX PI Penn SG, Hanzel DK, Chen W, Rank DR;

XX WPI; 2001-483446/52.

XX PT Single exon nucleic acid probes for analyzing gene expression in human

XX PT brains -

XX PS Example 4; SEQ ID NO: 25905; 650pp + Sequence Listing; English.

XX CC The present invention provides a number of single exon nucleic acid

CC probes which are derived from genomic sequences expressed in the human

CC brain. They can be used to measure gene expression in brain cell samples,

CC which may enable the diagnosis and improved treatment of nervous system

CC diseases such as Alzheimer's disease, multiple sclerosis, schizophrenia,

CC epilepsy and cancers. The present sequence is a protein encoded by one of

CC the probes of the invention.

XX SQ Sequence 51 AA;

XX Query Match 59.5%; Score 25; DB 22; Length 51;

XX Best Local Similarity 71.4%; Pred. No. 2.2e+02;

XX Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 3 LMAMNII 9

Db 17 LMAMNTL 23

RESULT 40

AAM66186

ID AAM66186 standard; Protein; 51 AA.

XX AC AAM66186;

XX 06-NOV-2001 (first entry)

XX DT

XX

Human bone marrow expressed probe encoded protein SEQ ID NO: 26492.

Human; bone marrow expressed exon; gene expression analysis; probe; microarray; cancer; leukaemia; lymphoma; myeloma.

Homo sapiens.

WO200157276-A2.

09-AUG-2001.

30-JAN-2001; 2001WO-US00668.

04-FEB-2000; 2000US-0180312.

26-MAY-2000; 2000US-0207456.

30-JUN-2000; 2000US-0608408.

03-AUG-2000; 2000US-0632366.

21-SEP-2000; 2000US-0234887.

27-SEP-2000; 2000US-0236359.

04-OCT-2000; 2000GB-0024263.

(MOLE-) MOLECULAR DYNAMICS INC.

Penn SG, Hanzel DK, Chen W, Rank DR;

WPI; 2001-488900/53.

Human genome-derived single exon nucleic acid probes useful for analyzing gene expression in human bone marrow -

Example 4; SEQ ID NO: 26492; 658pp + Sequence Listing; English.

The present invention provides a number of single exon nucleic acid probes which are derived from genomic sequences expressed in the human bone marrow. They can be used to measure gene expression in bone marrow samples, which may enable the improved diagnosis and treatment of cancers such as lymphoma, leukaemia and myeloma. The present sequence is a protein encoded by one of the probes of the invention.

Sequence 51 AA;

Query Match 59.5%; Score 25; DB 22; Length 51;  
Best Local Similarity 71.4%; Pred. NO. 2.2e+02;  
Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 3 LMAMNII 9  
| | | | | :  
Db 17 LMAMNTL 23

RESULT 41  
AAM14057  
ID AAM14057 standard; Protein; 51 AA.  
AC AAM14057;  
XX  
XX 12-OCT-2001 (first entry)  
XX  
XX Peptide #491 encoded by probe for measuring cervical gene expression.  
DE Probe; human; microarray; gene expression; cervical epithelial cell;  
KW cervical cancer.  
XX Homo sapiens.  
OS  
XX WO200157278-A2.  
PN  
XX  
XX 09-AUG-2001.  
PD  
XX  
XX 30-JAN-2001; 2001WO-US00670.  
PF  
XX  
XX 04-FEB-2000; 2000US-0180312.  
PR  
XX 26-MAY-2000; 2000US-0207456.  
PR

DR WPI; 2001-48897/53.  
 XX Human genome-derived single exon nucleic acid probes useful for  
 PT analyzing gene expression in human placenta -  
 XX  
 XX  
 PS Claim 27; SEQ ID No 26732; 654pp; English.  
 CC  
 CC The present invention relates to single exon nucleic acid probes (SENP;  
 CC see AAI31315-AA157546). The present sequence is a peptide encoded by one  
 CC such probe. The probes are useful for producing a microarray for  
 CC predicting, measuring and displaying gene expression in samples derived  
 CC from human placenta. The probes are useful for antenatal diagnosis of  
 CC human genetic disorders.  
 XX  
 XX  
 SQ Sequence 51 AA;  
 Query Match 59.5%; Score 25; DB 22; Length 51;  
 Best Local Similarity 71.4%; Pred. No. 2.2e+02;  
 Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;  
 QY 3 LMAMNII 9  
 Db 17 LMAMNTL 23  
 RESULT 43  
 AAM01797  
 ID AAM01797 standard; Protein; 51 AA.  
 XX  
 AC AAM01797;  
 XX  
 XX 09-OCT-2001 (first entry)  
 XX Peptide #479 encoded by probe for measuring human breast gene expression.  
 DE  
 DE Probe; human; breast disease; breast cancer; development disorder;  
 KW inflammatory disease; proliferative breast disease; non-carcinoma tumour.  
 KW  
 XX Homo sapiens.  
 OS  
 XX  
 XX WO200157270-A2.  
 PN  
 XX 09-AUG-2001.  
 PD  
 XX 29-JAN-2001; 2001WO-US00661.  
 XX  
 XX 04-FEB-2000; 2000US-0180312.  
 PR  
 PR 26-MAY-2000; 2000US-0207456.  
 PR 30-JUN-2000; 2000US-0608408.  
 PR 03-AUG-2000; 2000US-0632366.  
 PR 21-SEP-2000; 2000US-0234687.  
 PR 27-SEP-2000; 2000US-0236359.  
 PR 04-OCT-2000; 2000GB-0024263.  
 XX  
 XX (MOLE-) MOLECULAR DYNAMICS INC.  
 PA  
 XX Penn SG, Hanzel DK, Chen W, Rank DR;  
 XX WPI; 2001-476286/51.  
 XX  
 XX Novel single exon nucleic acid probe used to measuring gene expression  
 PT in a human breast -  
 PT  
 XX  
 PS Claim 27; SEQ ID No 10537; 322pp; English.  
 XX  
 CC The present invention relates to novel single exon nucleic acid probes  
 CC (see AAI00010-AA110067). The present sequence is a peptide encoded by one  
 CC such probe. The probes are useful for measuring human gene expression in  
 CC a human breast sample, where the probe hybridises at high stringency to a  
 CC nucleic acid expressed in the human breast. The probes are useful for  
 CC predicting, diagnosing, grading, staging, monitoring and prognosing  
 CC diseases of the human breast, particularly those diseases with polygenic  
 CC aetiology. The diseases include: breast cancer, disorders of development,

CC inflammatory diseases of the breast, fibrocystic changes, proliferative  
 CC breast disease and non-carcinoma tumours.  
 CC Note: The sequence data for this patent did not form part of the printed  
 CC specification, but was obtained in electronic format directly from WIPO  
 CC at ftp.wipo.int/pub/published\_pct\_sequences.  
 XX  
 SQ Sequence 51 AA;  
 Query Match 59.5%; Score 25; DB 22; Length 51;  
 Best Local Similarity 71.4%; Pred. No. 2.2e+02;  
 Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;  
 QY 3 LMAMNII 9  
 Db 17 LMAMNTL 23  
 RESULT 44  
 ABG35836  
 ID ABG35836 standard; Peptide; 51 AA.  
 XX  
 AC ABG35836;  
 XX  
 XX 19-AUG-2002 (first entry)  
 DT  
 XX Human peptide encoded by genome-derived single exon probe SEQ ID 25501.  
 DE  
 XX Human; single exon probe; asthma; lung cancer; COPD; ILD;  
 KW chronic obstructive pulmonary disease; interstitial lung disease;  
 KW familial idiopathic pulmonary fibrosis; neurofibromatosis;  
 KW tuberous sclerosis; Gaucher's disease; Niemann-Pick disease;  
 KW Hermansky-Pudlak syndrome; sarcoidosis; pulmonary haemosiderosis;  
 KW pulmonary histiocytosis; lymphangioleiomyomatosis; Karagenen syndrome;  
 KW pulmonary alveolar proteinosis; fibrocystic pulmonary dysplasia;  
 KW primary ciliary dyskinesia; pulmonary hypertension;  
 KW hyaline membrane disease.  
 KW  
 XX Homo sapiens.  
 OS  
 XX  
 XX WO200186003-A2.  
 PN  
 XX 15-NOV-2001.  
 PD  
 XX 30-JAN-2001; 2001WO-US00665.  
 XX  
 XX 04-FEB-2000; 2000US-180312P.  
 PR  
 PR 26-MAY-2000; 2000US-207456P.  
 PR 30-JUN-2000; 2000US-0608408.  
 PR 03-AUG-2000; 2000US-0632366.  
 PR 21-SEP-2000; 2000US-234687P.  
 PR 27-SEP-2000; 2000US-236359P.  
 PR 04-OCT-2000; 2000GB-0024263.  
 XX  
 XX (MOLE-) MOLECULAR DYNAMICS INC.  
 PA  
 XX Penn SG, Hanzel DK, Chen W, Rank DR;  
 XX WPI; 2002-114183/15.  
 XX  
 XX Spatially-addressable set of single exon nucleic acid probes, used to  
 PT measure gene expression in human lung samples -  
 PT  
 XX  
 PS Claim 27; SEQ ID No 25501; 634pp; English.  
 XX  
 CC The invention relates to a spatially-addressable set of single exon  
 CC nucleic acid probes for measuring gene expression in a sample derived  
 CC from human lung comprising single exon nucleic acid probes having one of  
 CC 12614 nucleic acid sequences mentioned in the specification, or their  
 CC complements or the 12387 open reading frames derived from the 12614  
 CC probes. Also included are a microarray comprising the novel set of  
 CC probes; the novel set of probes which hybridise at high stringency to a  
 CC nucleic acid expressed in the human lung; measuring gene expression in a  
 CC sample derived from human lung, comprising (a) contacting the array with

CC a collection of detectably labeled nucleic acids derived from human lung  
CC mRNA, and (b) measuring the label detectably bound to each probe of  
CC the array; identifying exons in a eukaryotic genome, comprising  
CC (a) algorithmically predicting at least one exon from genomic sequences  
CC of the eukaryote; and (b) detecting specific hybridisation of detectably  
CC labeled nucleic acids from eukaryote lung mRNA, to a single exon probe,  
CC having a fragment identical to the predicted exon, the probe is included  
CC in the above mentioned microarray; assigning exons to a single gene,  
CC comprising (a) identifying exons from genomic sequence by the method  
CC above and (b) measuring the expression of each of the exons in several  
CC tissues and/or cell types using hybridisation to a single exon  
CC microarrays having a probe with the exon, where a common pattern of  
CC expression of the exons in the tissues and/or cell types indicates that  
CC the exons should be assigned to a single gene; a peptide comprising one  
CC of 1201 sequences, mentioned in the specification, or encoded by the  
CC probes/open reading frames (ORF). The probes are used for gene  
CC expression analysis, and for identifying exons in a gene, particularly  
CC using human lung derived mRNA and for the study of lung diseases  
CC such as asthma, lung cancer, chronic obstructive pulmonary disease  
CC (COPD), interstitial lung disease (ILD), familial idiopathic pulmonary  
CC fibrosis, neurofibromatosis, tuberous sclerosis, Gaucher's disease,  
CC Niemann-Pick disease, Hermansky-Pudlak syndrome, sarcoidosis, pulmonary  
CC haemorrhoidosis, pulmonary histiocytosis, lymphangioleiomyomatosis,  
CC pulmonary alveolar proteinosis, Karagenen syndrome, fibrocystic  
CC pulmonary dysplasia, primary ciliary dyskinesia, pulmonary hypertension  
CC and hyaline membrane disease. The present sequence is a peptide/protein  
CC encoded by a single exon probe of the invention.  
CC Note: The sequence data for this patent did not form part  
CC of the printed specification, but was obtained in electronic  
CC format directly from WIPO at  
CC ftp.wipo.int/pub/published\_pct\_sequences.  
XX SQ Sequence 51 AA;  
Query Match 59.5%; Score 25; DB 23; Length 51;  
Best Local Similarity 71.4%; Pred. No. 2.2e+02;  
Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;  
QY 3 LMAMNII 9  
Db 17 LMAMNTL 23  
RESULT 45  
AAU14232  
ID AAU14232 standard; Protein; 52 AA.  
AC AAU14232;  
XX 24-OCT-2001 (first entry)  
DE Human novel protein #103.  
XX Human; novel protein; Antianaemic; osteopathic; antiinflammatory;  
KW immunomodulatory; cytostatic; neuroprotective; vulnerary; nootropic;  
KW anticonvulsant; antiarthritic; cerebroprotective; antifungal; antiviral;  
KW antibacterial; antiallergic; dermatological; haemostatic; antiasthmatic;  
KW thrombolytic; immunogen; antibody; gene therapy; neurological disorder;  
KW Parkinson's disease; inflammatory disorder; cancer; asthma; osteoporosis;  
KW tissue regeneration; immune disorder.  
XX Homo sapiens.  
OS WO200155437-A2.  
XX 02-AUG-2001.  
XX 25-JAN-2001; 2001WO-US02623.  
XX 25-JAN-2000; 2000US-0491404.  
XX (HYSE-) HYSEQ INC.  
XX

PI Tang YT, Liu C, Drmanac RT;  
XX WPI; 2001-451939/48.  
DR N-PSDB; AAS22537.  
XX Isolated polypeptides useful for treating anti-inflammatory diseases,  
PT nervous system disorders, and for regenerating bone and cartilage -  
XX Example 4; Page 581; 894pp; English.  
PS The invention relates to polynucleotides encoding novel human  
XX proteins or their active domains. The polypeptides, polynucleotides and  
CC antibodies raised against the polypeptides are used in a method of  
CC treatment of a mammal and prevention of disorders caused by the aberrant  
CC protein expression or activity. The polypeptides can be used as  
CC molecular weight markers, food supplements, and in antibody production.  
CC The polypeptides are used to identify compounds which bind to the  
CC polypeptides. Polynucleotides of the invention are used as probes and  
CC primers, for sequencing, for chromosome or gene mapping, in the  
CC production of recombinant proteins, and in generating anti-sense DNA or  
CC RNA and in gene therapy. Polypeptides of the invention can be used to  
CC target drugs to a tumour, in assays to determine biological activity, to  
CC raise antibodies/elicit an immune response, to determine quantitative  
CC protein levels, as tissue markers, and to isolate receptors or ligands.  
CC Polypeptides of the invention may also be useful in treating platelet  
CC disorders, stem cell disorders, regenerating bone, cartilage, tendon,  
CC ligament and/or nerve tissue, wound healing, treating burns, promoting  
CC the proliferation, differentiation and survival of stem cells, as a  
CC contraceptive, treating osteoporosis and osteoarthritis, anaemia,  
CC Alzheimer's, Parkinson's and Huntington's diseases, amyotrophic lateral  
CC sclerosis, stroke, immune deficiencies resulting from bacterial, viral or  
CC fungal infection or from autoimmunity, cancer, allergy, asthma,  
CC graft-versus-host disease, eczema, haemophilia, thrombosis,  
CC anti-inflammatory diseases, nervous system disorders, and infection.  
CC The present sequence represents a protein of the invention.  
XX SQ Sequence 52 AA;  
Query Match 59.5%; Score 25; DB 22; Length 52;  
Best Local Similarity 83.3%; Pred. No. 2.2e+02;  
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;  
QY 4 MAMNII 9  
Db 1 MALNII 6  
Search completed: February 17, 2004, 10:53:40  
Job time : 18.1089 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: February 17, 2004, 10:53:48 ; Search time 13.1881 Seconds  
(without alignments)  
142.889 Million cell updates/sec

Title: US-09-900-147-2  
Perfect score: 42  
Sequence: 1 NVLMANNII 9

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 801455 seqs, 209382283 residues

Total number of hits satisfying chosen parameters: 318354

Minimum DB seq length: 0  
Maximum DB seq length: 100

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Published Applications AA:  
1: /cgn2\_6/ptodata/1/pubpaa/US07\_PUBCOMB.pep.\*  
2: /cgn2\_6/ptodata/1/pubpaa/PCT\_NEW\_PUB.pep.\*  
3: /cgn2\_6/ptodata/1/pubpaa/US06\_NEW\_PUB.pep.\*  
4: /cgn2\_6/ptodata/1/pubpaa/US06\_PUBCOMB.pep.\*  
5: /cgn2\_6/ptodata/1/pubpaa/US07\_NEW\_PUB.pep.\*  
6: /cgn2\_6/ptodata/1/pubpaa/PCTUS\_PUBCOMB.pep.\*  
7: /cgn2\_6/ptodata/1/pubpaa/US08\_NEW\_PUB.pep.\*  
8: /cgn2\_6/ptodata/1/pubpaa/US08\_PUBCOMB.pep.\*  
9: /cgn2\_6/ptodata/1/pubpaa/US09A\_PUBCOMB.pep.\*  
10: /cgn2\_6/ptodata/1/pubpaa/US09B\_PUBCOMB.pep.\*  
11: /cgn2\_6/ptodata/1/pubpaa/US09C\_PUBCOMB.pep.\*  
12: /cgn2\_6/ptodata/1/pubpaa/US09\_NEW\_PUB.pep.\*  
13: /cgn2\_6/ptodata/1/pubpaa/US10A\_PUBCOMB.pep.\*  
14: /cgn2\_6/ptodata/1/pubpaa/US10B\_PUBCOMB.pep.\*  
15: /cgn2\_6/ptodata/1/pubpaa/US10C\_PUBCOMB.pep.\*  
16: /cgn2\_6/ptodata/1/pubpaa/US10\_NEW\_PUB.pep.\*  
17: /cgn2\_6/ptodata/1/pubpaa/US60\_NEW\_PUB.pep.\*  
18: /cgn2\_6/ptodata/1/pubpaa/US60\_PUBCOMB.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	42	100.0	9	10	US-09-900-147-2
2	42	100.0	16	10	US-09-900-147-5
3	42	100.0	19	10	US-09-900-147-3
4	42	100.0	19	10	US-09-900-147-16
5	42	100.0	20	10	US-09-900-147-4
6	42	100.0	30	10	US-09-900-147-6
7	42	100.0	37	10	US-09-900-147-1
8	42	100.0	74	15	US-10-214-188-10
9	38	90.5	19	10	US-09-900-147-15
10	36	85.7	19	10	US-09-900-147-17
11	34	81.0	14	10	US-09-900-147-11
12	31	73.8	67	9	US-09-864-761-43263
13	29	69.0	88	9	US-09-939-980-513
14	26	61.9	55	15	US-10-106-698-6066
15	25	59.5	25	9	US-09-864-761-48875

16	25	59.5	35	11	US-09-892-877-228	Sequence 228, App
17	25	59.5	35	11	US-09-948-783-230	Sequence 230, App
18	25	59.5	50	12	US-10-029-386-31645	Sequence 31645, A
19	25	59.5	51	9	US-09-864-761-33773	Sequence 33773, A
20	25	59.5	52	12	US-10-291-265-339	Sequence 339, App
21	25	59.5	52	12	US-10-291-265-811	Sequence 811, App
22	25	59.5	56	14	US-10-044-359-8	Sequence 8, Appli
23	25	59.5	66	9	US-09-864-761-47543	Sequence 47543, A
24	25	59.5	66	12	US-09-933-767-1030	Sequence 1030, Ap
25	25	59.5	66	15	US-10-023-282-1030	Sequence 1030, Ap
26	25	59.5	69	12	US-09-864-408A-5276	Sequence 5276, Ap
27	25	59.5	78	9	US-09-864-761-33623	Sequence 33623, A
28	25	59.5	84	12	US-09-933-767-1032	Sequence 1032, Ap
29	25	59.5	84	15	US-10-023-282-1032	Sequence 1032, Ap
30	25	59.5	88	12	US-10-029-386-33793	Sequence 33793, A
31	24.5	58.3	64	9	US-09-864-761-44843	Sequence 44843, A
32	24	57.1	8	12	US-10-376-121A-66	Sequence 66, Appl
33	24	57.1	25	9	US-09-864-761-46815	Sequence 46815, A
34	24	57.1	25	15	US-10-120-604-112	Sequence 112, App
35	24	57.1	28	9	US-09-864-761-40663	Sequence 40663, A
36	24	57.1	29	12	US-10-387-977-27	Sequence 27, Appl
37	24	57.1	37	15	US-10-224-260-6	Sequence 6, Appli
38	24	57.1	41	9	US-09-864-761-40202	Sequence 40202, A
39	24	57.1	44	12	US-10-321-857-54	Sequence 54, Appl
40	24	57.1	44	12	US-10-318-675-54	Sequence 54, Appl
41	24	57.1	52	12	US-10-231-417-202	Sequence 202, App
42	24	57.1	57	9	US-09-864-761-36764	Sequence 36764, A
43	24	57.1	58	14	US-10-044-359-2	Sequence 2, Appli
44	24	57.1	59	14	US-10-044-359-10	Sequence 10, Appl
45	24	57.1	62	12	US-09-864-408A-72	Sequence 72, Appl

ALIGNMENTS

RESULT 1  
US-09-900-147-2  
; Sequence 2, Application US/09900147  
; Patent No. US20020103121A1  
; GENERAL INFORMATION:  
; APPLICANT: La Thangue, Nicholas B  
; TITLE OF INVENTION: Peptide antagonists of DP transcription factors  
; FILE REFERENCE: 620-67  
; CURRENT FILING DATE: 2001-07-09  
; PRIOR FILING DATE: 2001-07-09  
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 09/308,935  
; PRIOR FILING DATE: EARLIER FILING DATE: 1999-05-27  
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: GB 9626589.7  
; NUMBER OF SEQ ID NOS: 18  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 2  
; LENGTH: 9  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Description of Artificial Sequence: Synthetic  
US-09-900-147-2

Query Match 100.0%; Score 42; DB 10; Length 9;  
Best Local Similarity 100.0%; Pred. No. 7e+05;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 NVLMANNII 9  
| | | | |  
Db 1 NVLMANNII 9

RESULT 2  
US-09-900-147-5  
; Sequence 5, Application US/09900147  
; Patent No. US20020103121A1

```

; GENERAL INFORMATION:
; APPLICANT: La Thangue, Nicholas B
; APPLICANT: Bandara, Lasantha R
; TITLE OF INVENTION: Peptide antagonists of DP transcription factors
; FILE REFERENCE: 620-67
; CURRENT APPLICATION NUMBER: US/09/900,147
; CURRENT FILING DATE: 2001-07-09
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 09/308,935
; PRIOR FILING DATE: EARLIER FILING DATE: 1999-05-27
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: GB 9626589.7
; PRIOR FILING DATE: EARLIER FILING DATE: 1996-12-20
; NUMBER OF SEQ ID NOS: 18
; SOFTWARE: Patentin Ver. 2.1
; SEQ ID NO 5
; LENGTH: 16
; TYPE: PPT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Synthetic
US-09-900-147-5

Query Match          100.0%; Score 42; DB 10; Length 16;
Best Local Similarity 100.0%; Pred. No. 0.057;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 NVLMAMNII 9
DB      7 NVLMAMNII 15

RESULT 3
US-09-900-147-3
; Sequence 3, Application US/09900147
; Patent No. US20020103121A1
; GENERAL INFORMATION:
; APPLICANT: La Thangue, Nicholas B
; APPLICANT: Bandara, Lasantha R
; TITLE OF INVENTION: Peptide antagonists of DP transcription factors
; FILE REFERENCE: 620-67
; CURRENT APPLICATION NUMBER: US/09/900,147
; CURRENT FILING DATE: 2001-07-09
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 09/308,935
; PRIOR FILING DATE: EARLIER FILING DATE: 1999-05-27
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: GB 9626589.7
; PRIOR FILING DATE: EARLIER FILING DATE: 1996-12-20
; NUMBER OF SEQ ID NOS: 18
; SOFTWARE: Patentin Ver. 2.1
; SEQ ID NO 3
; LENGTH: 19
; TYPE: PPT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Synthetic
US-09-900-147-3

Query Match          100.0%; Score 42; DB 10; Length 19;
Best Local Similarity 100.0%; Pred. No. 0.068;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 NVLMAMNII 9
DB      9 NVLMAMNII 17

RESULT 4
US-09-900-147-16
; Sequence 16, Application US/09900147
; Patent No. US20020103121A1
; GENERAL INFORMATION:
; APPLICANT: La Thangue, Nicholas B
; APPLICANT: Bandara, Lasantha R
; TITLE OF INVENTION: Peptide antagonists of DP transcription factors

```



; PRIOR FILING DATE: EARLIER FILING DATE: 1996-12-20  
; NUMBER OF SEQ ID NOS: 18  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 6  
; LENGTH: 30  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Description of Artificial Sequence: Synthetic  
US-09-900-147-6

Query Match 100.0%; Score 42; DB 10; Length 30;  
Best Local Similarity 100.0%; Pred. No. 0.11; Indels 0; Gaps 0;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 NVLMANNII 9  
Db 5 NVLMANNII 13

RESULT 7  
US-09-900-147-1  
; Sequence 1, Application US/09900147  
; Patent No. US20020103121A1  
; GENERAL INFORMATION:  
; APPLICANT: La Thangue, Nicholas B  
; APPLICANT: Bandara, Lasantha R  
; TITLE OF INVENTION: Peptide antagonists of DP transcription factors  
; FILE REFERENCE: 620-67  
; CURRENT APPLICATION NUMBER: US/09/900,147  
; CURRENT FILING DATE: 2001-07-09  
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 09/308,935  
; PRIOR FILING DATE: EARLIER FILING DATE: 1999-05-27  
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: GB 9626589.7  
; PRIOR FILING DATE: EARLIER FILING DATE: 1996-12-20  
; NUMBER OF SEQ ID NOS: 18  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 1  
; LENGTH: 37  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Description of Artificial Sequence: Synthetic  
US-09-900-147-1

Query Match 100.0%; Score 42; DB 10; Length 37;  
Best Local Similarity 100.0%; Pred. No. 0.14; Indels 0; Gaps 0;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 NVLMANNII 9  
Db 12 NVLMANNII 20

RESULT 8  
US-10-214-188-10  
; Sequence 10, Application US/10214188  
; Publication No. US20030022260A1  
; GENERAL INFORMATION:  
; APPLICANT: LA THANGUE, NICHOLAS B.  
; BERNARDS, RENE  
; HIJMAN, ELEANORE M.  
; TITLE OF INVENTION: TRANSCRIPTION FACTOR E2F-5  
; NUMBER OF SEQUENCES: 25  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: NIXON & VANDERHYE P. C.  
; STREET: 1100 NORTH GLEBE ROAD  
; CITY: ARLINGTON  
; STATE: VIRGINIA  
; COUNTRY: U.S.A.  
; ZIP: 22201-4714  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/10/214,188  
; FILING DATE: 08-Aug-2002  
; CLASSIFICATION: <Unknown>  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US/08/894,139  
; FILING DATE: 13-AUG-1997  
; ATTORNEY/AGENT INFORMATION:  
; NAME: WILSON, MARY J.  
; REGISTRATION NUMBER: 32,955  
; REFERENCE/DOCKET NUMBER: 620-22  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (703) 816-4000  
; TELEFAX: (703) 816-4100  
; INFORMATION FOR SEQ ID NO: 10:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 74 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: <Unknown>  
; TOPOLOGY: linear  
; MOLECULE TYPE: peptide  
; SEQUENCE DESCRIPTION: SEQ ID NO: 10:  
US-10-214-188-10

Query Match 100.0%; Score 42; DB 15; Length 74;  
Best Local Similarity 100.0%; Pred. No. 0.29; Indels 0; Gaps 0;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 NVLMANNII 9  
Db 54 NVLMANNII 62

RESULT 9  
US-09-900-147-15  
; Sequence 15, Application US/09900147  
; Patent No. US20020103121A1  
; GENERAL INFORMATION:  
; APPLICANT: La Thangue, Nicholas B  
; APPLICANT: Bandara, Lasantha R  
; TITLE OF INVENTION: Peptide antagonists of DP transcription factors  
; FILE REFERENCE: 620-67  
; CURRENT APPLICATION NUMBER: US/09/900,147  
; CURRENT FILING DATE: 2001-07-09  
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 09/308,935  
; PRIOR FILING DATE: EARLIER FILING DATE: 1999-05-27  
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: GB 9626589.7  
; PRIOR FILING DATE: EARLIER FILING DATE: 1996-12-20  
; NUMBER OF SEQ ID NOS: 18  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 15  
; LENGTH: 19  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Description of Artificial Sequence: Mutant peptide  
US-09-900-147-15

Query Match 90.5%; Score 38; DB 10; Length 19;  
Best Local Similarity 88.9%; Pred. No. 0.42; Indels 1; Gaps 0;  
Matches 8; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 NVLMANNII 9  
Db 9 NALMANNII 17

RESULT 10  
US-09-900-147-17  
; Sequence 17, Application US/09900147

Tue Feb 17 11:55:47 2004

us-09-900-147-2.rapb

```

; Patent No. US20020103121A1
; GENERAL INFORMATION:
; APPLICANT: La Thangue, Nicholas B
; APPLICANT: Bandara, Lasantha R
; TITLE OF INVENTION: Peptide antagonists of DP transcription factors
; FILE REFERENCE: 620-67
; CURRENT APPLICATION NUMBER: US/09/900,147
; CURRENT FILING DATE: 2001-07-09
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 09/308,935
; PRIOR FILING DATE: EARLIER FILING DATE: 1999-05-27
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: GB 9626589.7
; PRIOR FILING DATE: EARLIER FILING DATE: 1996-12-20
; NUMBER OF SEQ ID NOS: 18
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 17
; LENGTH: 19
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Mutant peptide
US-09-900-147-17

Query Match 85.7%; Score 36; DB 10; Length 19;
Best Local Similarity 88.9%; Pred. No. 1;
Matches 8; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 NVLMAMNII 9
Db 9 NVLMAMNII 17

RESULT 11
US-09-900-147-11
; Sequence 11, Application US/09900147
; Patent No. US20020103121A1
; GENERAL INFORMATION:
; APPLICANT: La Thangue, Nicholas B
; APPLICANT: Bandara, Lasantha R
; TITLE OF INVENTION: Peptide antagonists of DP transcription factors
; FILE REFERENCE: 620-67
; CURRENT APPLICATION NUMBER: US/09/900,147
; CURRENT FILING DATE: 2001-07-09
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 09/308,935
; PRIOR FILING DATE: EARLIER FILING DATE: 1999-05-27
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: GB 9626589.7
; PRIOR FILING DATE: EARLIER FILING DATE: 1996-12-20
; NUMBER OF SEQ ID NOS: 18
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 11
; LENGTH: 14
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Synthetic.
US-09-900-147-11

Query Match 81.0%; Score 34; DB 10; Length 14;
Best Local Similarity 100.0%; Pred. No. 1.9;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 NVLMAMN 7
Db 8 NVLMAMN 14

RESULT 12
US-09-864-761-43263
; Sequence 43263, Application US/09864761
; Patent No. US20020048763A1
; GENERAL INFORMATION:
; APPLICANT: Penn, Sharon G.
; APPLICANT: Rank, David R.
; APPLICANT: Hanzel, David K.
; APPLICANT: Chen, Wensheng
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
; TITLE OF INVENTION: GENE EXPRESSION ANALYSIS BY MICROARRAY
; FILE REFERENCE: Aecmica-X-1
; CURRENT APPLICATION NUMBER: US/09/864,761
; CURRENT FILING DATE: 2001-05-23
; PRIOR APPLICATION NUMBER: US 60/180,312
; PRIOR FILING DATE: 2000-02-04
; PRIOR APPLICATION NUMBER: US 60/207,456
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: US 09/632,366
; PRIOR FILING DATE: 2000-08-03
; PRIOR APPLICATION NUMBER: GB 24263.6
; PRIOR FILING DATE: 2000-10-04
; PRIOR APPLICATION NUMBER: US 60/236,359
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: PCT/US01/00666
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00667
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00664
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00669
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00665
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00668
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00663
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00662
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00661
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00670
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: US 60/234,687
; PRIOR FILING DATE: 2000-09-21
; PRIOR APPLICATION NUMBER: US 09/608,408
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: US 09/774,203
; PRIOR FILING DATE: 2001-01-29
; NUMBER OF SEQ ID NOS: 49117
; SOFTWARE: Annomax Sequence Listing Engine vers. 1.1
; SEQ ID NO 43263
; LENGTH: 67
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: MAP TO AL031301.1
; OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 1.6
US-09-864-761-43263

Query Match 73.8%; Score 31; DB 9; Length 67;
Best Local Similarity 55.6%; Pred. No. 39;
Matches 5; Conservative 4; Mismatches 0; Indels 0; Gaps 0;

Qy 1 NVLMAMNII 9
Db 43 NVLMAMNII 51

RESULT 13
US-09-939-980-513
; Sequence 513, Application US/09939980
; Patent No. US2002008234A1
; GENERAL INFORMATION:
; APPLICANT: Black, Michael
; APPLICANT: Burnham, Martin
; APPLICANT: Hodgson, John
; APPLICANT: Knowles, David
; APPLICANT: Lonetto, Michael
; APPLICANT: Nicholas, Richard

```

```
;
; Pratt, Julie
; Reichard, Richard
; Rosenberg, Martin
; Ward, Judith
;
; TITLE OF INVENTION: No. US200200822341el Prokaryotic Polynucleotides,
; Polypeptides and Their Uses
;
; NUMBER OF SEQUENCES: 534
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: SmithKline Beecham Corporation
; STREET: 709 Swedeland Road
; CITY: King of Prussia
; STATE: PA
; COUNTRY: USA
; ZIP: 19406-0939
;
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; FILING APPLICATION NUMBER: US/09/939,980
; FILING DATE: 27-Aug-2001
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/936,165
; FILING DATE: <Unknown>
; ATTORNEY/AGENT INFORMATION:
; NAME: Gimmi, Edward R
; REGISTRATION NUMBER: 38,891
; REFERENCE/DOCKET NUMBER: P50549
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 610-270-4478
; TELEFAX: 610-270-5090
; TELEX: <Unknown>
;
; INFORMATION FOR SEQ ID NO: 513:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 88 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: Protein
; SEQUENCE DESCRIPTION: SEQ ID NO: 513:
;
; us-09-939-980-513
```

```
Query Match 69.0%; Score 29; DB 9; Length 88;
Best Local Similarity 66.7%; Pred. No. 1.3e+02;
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;
```

```
Qy 1 NLMAMNII 9
|||:||||
Db 55 NLVLPINII 63
```

```
RESULT 14
US-10-106-698-6066
; Sequence 6066, Application US/10106698
; Publication No. US20030109690A1
; GENERAL INFORMATION:
; APPLICANT: Ruben et al.
; TITLE OF INVENTION: Colon and Colon Cancer Associated Polynucleotides and Polypeptide
; FILE REFERENCE: PA005P1
; CURRENT APPLICATION NUMBER: US/10/106,698
; CURRENT FILING DATE: 2002-03-27
; PRIOR APPLICATION NUMBER: PCT/US00/26524
; PRIOR FILING DATE: 2000-09-28
; PRIOR APPLICATION NUMBER: US 60/157,137
; PRIOR FILING DATE: 1999-09-29
; PRIOR APPLICATION NUMBER: US 60/163,280
; PRIOR FILING DATE: 1999-11-03
; NUMBER OF SEQ ID NOS: 8564
; SOFTWARE: PatentIn ver. 3.0
; SEQ ID NO 6066
; LENGTH: 55
```

```
;
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: MISC FEATURE
; LOCATION: (5)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; US-10-106-698-6066
;
; Query Match 61.9%; Score 26; DB 15; Length 55;
; Best Local Similarity 62.5%; Pred. No. 3.1e+02;
; Matches 5; Conservative 2; Mismatches 1; Indels 0; Gaps 0;
;
; Qy 2 VLMAMNII 9
; Db 45 VILAKNIV 52
; |||:||||
;
; RESULT 15
; US-09-864-761-48875
; Sequence 48875, Application US/09864761
; Patent No. US20020048763A1
; GENERAL INFORMATION:
; APPLICANT: Penn, Sharron G.
; APPLICANT: Rank, David R.
; APPLICANT: Hanzel, David K.
; APPLICANT: Chen, Wensheng
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
; FILE REFERENCE: Aeomica-X-1
; CURRENT APPLICATION NUMBER: US/09/864,761
; CURRENT FILING DATE: 2001-05-23
; PRIOR APPLICATION NUMBER: US 60/180,312
; PRIOR FILING DATE: 2000-02-04
; PRIOR APPLICATION NUMBER: US 60/207,456
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: US 09/632,366
; PRIOR FILING DATE: 2000-08-03
; PRIOR APPLICATION NUMBER: GB 24263.6
; PRIOR FILING DATE: 2000-10-04
; PRIOR APPLICATION NUMBER: US 60/236,359
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: PCT/US01/00666
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00667
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00664
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00669
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00665
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00668
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00663
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00662
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00661
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00670
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: US 60/234,687
; PRIOR FILING DATE: 2000-09-21
; PRIOR APPLICATION NUMBER: US 09/608,408
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: US 09/774,203
; PRIOR FILING DATE: 2001-01-29
; NUMBER OF SEQ ID NOS: 49117
; SOFTWARE: Annonmax Sequence Listing Engine vers. 1.1
; SEQ ID NO 48875
; LENGTH: 25
; TYPE: PRT
; ORGANISM: Homo sapiens
```

```

; FEATURE:
; OTHER INFORMATION: MAP TO AC019181.2
; OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 2.7
; OTHER INFORMATION: EST_HUMAN HIT: BF677515.1, EVALUE 4.00e-07
US-09-864-761-48875

Query Match          59.5%; Score 25; DB 9; Length 25;
Best Local Similarity 33.3%; Pred. No. 2.le+02;
Matches 3; Conservative 5; Mismatches 1; Indels 0; Gaps 0;

Qy 1 NVLMAMNII 9
Db 15 NILMSIQLL 23

RESULT 16
US-09-892-877-228
; Sequence 228, Application US/09892877
; Publication No. US20030077809A1
; GENERAL INFORMATION:
; APPLICANT: Ruben et. al.
; TITLE OF INVENTION: 97 Human secreted proteins
; FILE REFERENCE: P2028P1
; CURRENT APPLICATION NUMBER: US/09/892,877
; CURRENT FILING DATE: 2001-06-28
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: US/09/437,658
; PRIOR FILING DATE: EARLIER FILING DATE: 1999-11-10
; NUMBER OF SEQ ID NOS: 461
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 228
; LENGTH: 35
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-892-877-228

Query Match          59.5%; Score 25; DB 11; Length 35;
Best Local Similarity 83.3%; Pred. No. 3e+02;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 4 MAMNII 9
Db 11 MALNII 16

RESULT 17
US-09-948-783-230
; Sequence 230, Application US/09948783
; Publication No. US20030100051A1
; GENERAL INFORMATION:
; APPLICANT: Ruben et. al.
; TITLE OF INVENTION: 97 Human secreted proteins
; FILE REFERENCE: P2028P2
; CURRENT APPLICATION NUMBER: US/09/948,783
; CURRENT FILING DATE: 2001-09-10
; PRIOR APPLICATION NUMBER: 60/231,846
; PRIOR FILING DATE: 2000-09-11
; PRIOR APPLICATION NUMBER: 09/892,877
; PRIOR FILING DATE: 2001-06-28
; PRIOR APPLICATION NUMBER: 09/437,658
; PRIOR FILING DATE: 1999-11-10
; PRIOR APPLICATION NUMBER: PCT/US99/09847
; PRIOR FILING DATE: 1999-05-06
; PRIOR APPLICATION NUMBER: 60/085,094
; PRIOR FILING DATE: 1998-05-12
; PRIOR APPLICATION NUMBER: 60/085,105
; PRIOR FILING DATE: 1998-05-12
; PRIOR APPLICATION NUMBER: 60/085,180
; PRIOR FILING DATE: 1998-05-12
; PRIOR APPLICATION NUMBER: 60/085,927
; PRIOR FILING DATE: 1998-05-18
; PRIOR APPLICATION NUMBER: 60/085,906
```

```

; PRIOR FILING DATE: 1998-05-18
; PRIOR APPLICATION NUMBER: 60/085,924
; PRIOR FILING DATE: 1998-05-18
; PRIOR APPLICATION NUMBER: 60/085,922
; PRIOR FILING DATE: 1998-05-18
; PRIOR APPLICATION NUMBER: 60/085,921
; PRIOR FILING DATE: 1998-05-18
; PRIOR APPLICATION NUMBER: 60/085,923
; PRIOR FILING DATE: 1998-05-18
; PRIOR APPLICATION NUMBER: 60/085,925
; PRIOR FILING DATE: 1998-05-18
; PRIOR APPLICATION NUMBER: 60/085,928
; PRIOR FILING DATE: 1998-05-18
; PRIOR APPLICATION NUMBER: 60/085,920
; NUMBER OF SEQ ID NOS: 465
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 230
; LENGTH: 35
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-948-783-230

Query Match          59.5%; Score 25; DB 11; Length 35;
Best Local Similarity 83.3%; Pred. No. 3e+02;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 4 MAMNII 9
Db 11 MALNII 16

RESULT 18
US-10-029-386-31645
; Sequence 31645, Application US/10029386
; Publication No. US20030194704A1
; GENERAL INFORMATION:
; APPLICANT: Penn, Sharron G.
; APPLICANT: Rank, David R.
; APPLICANT: Hanzel, David K.
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR G
; TITLE OF INVENTION: EXPRESSION ANALYSIS TWO
; FILE REFERENCE: AEOmica-X-2
; CURRENT APPLICATION NUMBER: US/10/029,386
; CURRENT FILING DATE: 2001-12-20
; NUMBER OF SEQ ID NOS: 34288
; SOFTWARE: Anomax Sequence Listing Engine vers. 1.1
; SEQ ID NO 31645
; LENGTH: 50
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: MAP TO AC005680.1
; OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 1.3
; OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 1.6
; OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 2.3
; OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 1.5
; OTHER INFORMATION: SWISSPROT HIT: P52746, EVALUE 6.30e+00
US-10-029-386-31645

Query Match          59.5%; Score 25; DB 12; Length 50;
Best Local Similarity 44.4%; Pred. No. 4.4e+02;
Matches 4; Conservative 4; Mismatches 1; Indels 0; Gaps 0;

Qy 1 NVLMAMNII 9
Db 21 NILMSIGIL 29

RESULT 19
US-09-864-761-33773
; Sequence 33773, Application US/09864761
; Patent No. US20020048763A1
```

GENERAL INFORMATION:  
; APPLICANT: Penn, Sharron G.  
; APPLICANT: Rank, David R.  
; APPLICANT: Hanzel, David K.  
; APPLICANT: Chen, Wensheng  
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR  
; FILE REFERENCE: Aecolica-X-1  
; CURRENT FILING DATE: 2001-05-23  
; PRIOR APPLICATION NUMBER: US 09/864,761  
; PRIOR FILING DATE: 2000-02-04  
; PRIOR APPLICATION NUMBER: US 60/180,312  
; PRIOR FILING DATE: 2000-05-26  
; PRIOR APPLICATION NUMBER: US 09/632,366  
; PRIOR FILING DATE: 2000-08-03  
; PRIOR APPLICATION NUMBER: GB 24263.6  
; PRIOR FILING DATE: 2000-10-04  
; PRIOR APPLICATION NUMBER: US 60/236,359  
; PRIOR FILING DATE: 2000-09-27  
; PRIOR APPLICATION NUMBER: PCT/US01/00666  
; PRIOR FILING DATE: 2001-01-30  
; PRIOR APPLICATION NUMBER: PCT/US01/00667  
; PRIOR FILING DATE: 2001-01-30  
; PRIOR APPLICATION NUMBER: PCT/US01/00664  
; PRIOR FILING DATE: 2001-01-30  
; PRIOR APPLICATION NUMBER: PCT/US01/00669  
; PRIOR FILING DATE: 2001-01-30  
; PRIOR APPLICATION NUMBER: PCT/US01/00665  
; PRIOR FILING DATE: 2001-01-30  
; PRIOR APPLICATION NUMBER: PCT/US01/00668  
; PRIOR FILING DATE: 2001-01-30  
; PRIOR APPLICATION NUMBER: PCT/US01/00663  
; PRIOR FILING DATE: 2001-01-30  
; PRIOR APPLICATION NUMBER: PCT/US01/00662  
; PRIOR FILING DATE: 2001-01-30  
; PRIOR APPLICATION NUMBER: PCT/US01/00661  
; PRIOR FILING DATE: 2001-01-30  
; PRIOR APPLICATION NUMBER: PCT/US01/00670  
; PRIOR FILING DATE: 2001-01-30  
; PRIOR APPLICATION NUMBER: US 60/234,687  
; PRIOR FILING DATE: 2000-09-21  
; PRIOR APPLICATION NUMBER: US 09/608,408  
; PRIOR FILING DATE: 2000-06-30  
; PRIOR APPLICATION NUMBER: US 09/774,203  
; PRIOR FILING DATE: 2001-01-29  
; NUMBER OF SEQ ID NOS: 49117  
; SOFTWARE: Annomax Sequence Listing Engine vers. 1.1  
; SEQ ID NO 33773  
; LENGTH: 51  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
; FEATURE:  
; OTHER INFORMATION: MAP TO AL034406.1  
; OTHER INFORMATION: EXPRESSED IN BT474, SIGNAL = 8.8  
; OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 1.3  
; OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 2.6  
; OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 1.8  
; OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 3.4  
; OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 2.4  
; OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 1  
; OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 1.5  
; OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 4.5  
; OTHER INFORMATION: EXPRESSED IN HEL100, SIGNAL = 2.6  
; OTHER INFORMATION: SWISSPROT HIT: P28044, EVALUATE 7.80e+00

Query Match 59.5%; Score 25; DB 9; Length 51;  
Best Local Similarity 71.4%; Pred. No. 4.5e+02;  
Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 3 LMAMNII 9  
|||||;

Db 17 LMAMNTL 23

RESULT 20  
US-10-291-265-339  
; Sequence 339, Application US/10291265  
; Publication No. US20030232054A1  
; GENERAL INFORMATION:  
; APPLICANT: Hyseq, Inc.  
; APPLICANT: Tang et al  
; TITLE OF INVENTION: No. US20030232054A1el Nucleic Acids and Polypeptides  
; FILE REFERENCE: 21272-017 (785)  
; CURRENT APPLICATION NUMBER: US/10/291,265  
; CURRENT FILING DATE: 2000-01-25  
; PRIOR APPLICATION NUMBER: 09/491,404  
; PRIOR FILING DATE: 2000-01-25  
; PRIOR APPLICATION NUMBER: 09/617,746  
; PRIOR FILING DATE: 2000-07-17  
; PRIOR APPLICATION NUMBER: 09/631,451  
; PRIOR FILING DATE: 2000-08-03  
; PRIOR APPLICATION NUMBER: 09/633,870  
; PRIOR FILING DATE: 2000-09-15  
; NUMBER OF SEQ ID NOS: 944  
; SOFTWARE: FastSeq for Windows Version 3.0  
; SEQ ID NO 339  
; LENGTH: 52  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-10-291-265-339

Query Match 59.5%; Score 25; DB 12; Length 52;  
Best Local Similarity 83.3%; Pred. No. 4.6e+02;  
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 4 MAMNII 9  
||:||||  
Db 1 MALNII 6

RESULT 21  
US-10-291-265-811  
; Sequence 811, Application US/10291265  
; Publication No. US20030232054A1  
; GENERAL INFORMATION:  
; APPLICANT: Hyseq, Inc.  
; APPLICANT: Tang et al  
; TITLE OF INVENTION: No. US20030232054A1el Nucleic Acids and Polypeptides  
; FILE REFERENCE: 21272-017 (785)  
; CURRENT APPLICATION NUMBER: US/10/291,265  
; CURRENT FILING DATE: 2000-01-25  
; PRIOR APPLICATION NUMBER: 09/491,404  
; PRIOR FILING DATE: 2000-01-25  
; PRIOR APPLICATION NUMBER: 09/617,746  
; PRIOR FILING DATE: 2000-07-17  
; PRIOR APPLICATION NUMBER: 09/631,451  
; PRIOR FILING DATE: 2000-08-03  
; PRIOR APPLICATION NUMBER: 09/633,870  
; PRIOR FILING DATE: 2000-09-15  
; NUMBER OF SEQ ID NOS: 944  
; SOFTWARE: FastSeq for Windows Version 3.0  
; SEQ ID NO 811  
; LENGTH: 52  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-10-291-265-811

Query Match 59.5%; Score 25; DB 12; Length 52;  
Best Local Similarity 83.3%; Pred. No. 4.6e+02;  
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 4 MAMNII 9  
||:||||  
Db 1 MALNII 6

RESULT 22  
US-10-044-359-8  
; Sequence 8, Application US/10044359  
; Publication No. US20020160454A1  
; GENERAL INFORMATION:  
; APPLICANT: Heirman, Rafael  
; APPLICANT: Wong, James F.  
; APPLICANT: Lee, Jian-Ming  
; TITLE OF INVENTION: SCORPION TOXINS  
; FILE REFERENCE: BBI387 US NA  
; CURRENT APPLICATION NUMBER: US/10/044,359  
; CURRENT FILING DATE: 2002-01-11  
; PRIOR APPLICATION NUMBER: 09/599,416  
; PRIOR FILING DATE: 2000-06-22  
; PRIOR APPLICATION NUMBER: 60/140,227  
; PRIOR FILING DATE: 1999-06-22  
; NUMBER OF SEQ ID NOS: 28  
; SOFTWARE: Microsoft Office 97  
; SEQ ID NO 8  
; LENGTH: 56  
; TYPE: PRT  
; ORGANISM: Hottentotta judaica  
US-10-044-359-8

Query Match 59.5%; Score 25; DB 14; Length 56;  
Best Local Similarity 37.5%; Pred. No. 5e+02;  
Matches 3; Conservative 5; Mismatches 0; Indels 0; Gaps 0;

QY 2 VLMANNII 9  
DB 10 IVLANNVM 17

RESULT 23  
US-09-864-761-47543  
; Sequence 47543, Application US/09864761  
; Patent No. US20020048763A1  
; GENERAL INFORMATION:  
; APPLICANT: Penn, Sharron G.  
; APPLICANT: Rank, David R.  
; APPLICANT: Hanzel, David K.  
; APPLICANT: Chen, Wenheng  
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR  
; FILE REFERENCE: Aeomica-X-1  
; CURRENT APPLICATION NUMBER: US/09/864,761  
; CURRENT FILING DATE: 2001-05-23  
; PRIOR APPLICATION NUMBER: US 60/180,312  
; PRIOR FILING DATE: 2000-02-04  
; PRIOR APPLICATION NUMBER: US 60/207,456  
; PRIOR FILING DATE: 2000-05-26  
; PRIOR APPLICATION NUMBER: US 09/632,366  
; PRIOR FILING DATE: 2000-08-03  
; PRIOR APPLICATION NUMBER: GB 24263.6  
; PRIOR FILING DATE: 2000-10-04  
; PRIOR APPLICATION NUMBER: US 60/236,359  
; PRIOR FILING DATE: 2000-09-27  
; PRIOR APPLICATION NUMBER: PCT/US01/00666  
; PRIOR FILING DATE: 2001-01-30  
; PRIOR APPLICATION NUMBER: PCT/US01/00667  
; PRIOR FILING DATE: 2001-01-30  
; PRIOR APPLICATION NUMBER: PCT/US01/00664  
; PRIOR FILING DATE: 2001-01-30  
; PRIOR APPLICATION NUMBER: PCT/US01/00669  
; PRIOR FILING DATE: 2001-01-30  
; PRIOR APPLICATION NUMBER: PCT/US01/00665  
; PRIOR FILING DATE: 2001-01-30  
; PRIOR APPLICATION NUMBER: PCT/US01/00668  
; PRIOR FILING DATE: 2001-01-30  
; PRIOR APPLICATION NUMBER: PCT/US01/00663  
; PRIOR FILING DATE: 2001-01-30

; PRIOR APPLICATION NUMBER: PCT/US01/00662  
; PRIOR FILING DATE: 2001-01-30  
; PRIOR APPLICATION NUMBER: PCT/US01/00661  
; PRIOR FILING DATE: 2001-01-30  
; PRIOR APPLICATION NUMBER: PCT/US01/00670  
; PRIOR FILING DATE: 2001-01-30  
; PRIOR APPLICATION NUMBER: US 60/234,687  
; PRIOR FILING DATE: 2000-09-21  
; PRIOR APPLICATION NUMBER: US 09/608,408  
; PRIOR FILING DATE: 2000-06-30  
; PRIOR APPLICATION NUMBER: US 09/774,203  
; PRIOR FILING DATE: 2001-01-29  
; NUMBER OF SEQ ID NOS: 49117  
; SOFTWARE: Anomax Sequence Listing Engine vers. 1.1  
; SEQ ID NO 47543  
; LENGTH: 66  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
; FEATURE:  
; OTHER INFORMATION: MAP TO AF124731.2  
; OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 0.72  
; OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 1.6  
; OTHER INFORMATION: EXPRESSED IN BT474, SIGNAL = 0.61  
; OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 0.98  
; OTHER INFORMATION: SWISSPROT HIT: O14867, EVALUATE 1.00e-32  
; OTHER INFORMATION: EST\_HUMAN HIT: BF512320.1, EVALUATE 2.00e-31  
US-09-864-761-47543

Query Match 59.5%; Score 25; DB 9; Length 66;  
Best Local Similarity 57.1%; Pred. No. 5.9e+02;  
Matches 4; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 1 NVLMANN 7  
DB 7 NVLLSLN 13

RESULT 24  
US-09-933-767-1030  
; Sequence 1030, Application US/09933767  
; Publication No. US20030181692A1  
; GENERAL INFORMATION:  
; APPLICANT: Ni et al.  
; TITLE OF INVENTION: 207 Human Secreted Proteins  
; FILE REFERENCE: P2007P2  
; CURRENT APPLICATION NUMBER: US/09/933,767  
; CURRENT FILING DATE: 2001-08-22  
; PRIOR APPLICATION NUMBER: PCT/US01/05614  
; PRIOR FILING DATE: 2001-02-21  
; PRIOR APPLICATION NUMBER: 60/184,836  
; PRIOR FILING DATE: 2000-02-24  
; PRIOR APPLICATION NUMBER: 60/193,170  
; PRIOR FILING DATE: 2000-03-29  
; PRIOR APPLICATION NUMBER: 09/205,258  
; PRIOR FILING DATE: 1998-12-04  
; PRIOR APPLICATION NUMBER: PCT/US98/11422  
; PRIOR FILING DATE: 1998-06-04  
; PRIOR APPLICATION NUMBER: 60/048,885  
; PRIOR FILING DATE: 1997-06-06  
; PRIOR APPLICATION NUMBER: 60/049,375  
; PRIOR FILING DATE: 1997-06-06  
; PRIOR APPLICATION NUMBER: 60/048,881  
; PRIOR FILING DATE: 1997-06-06  
; PRIOR APPLICATION NUMBER: 60/048,880  
; PRIOR FILING DATE: 1997-06-06  
; PRIOR APPLICATION NUMBER: 60/048,896  
; PRIOR FILING DATE: 1997-06-06  
; PRIOR APPLICATION NUMBER: 60/049,020  
; PRIOR FILING DATE: 1997-06-06  
; PRIOR APPLICATION NUMBER: 60/048,876  
; PRIOR FILING DATE: 1997-06-06  
; PRIOR APPLICATION NUMBER: 60/048,895  
; PRIOR FILING DATE: 1997-06-06

; PRIOR APPLICATION NUMBER: 60/048,884  
; PRIOR FILING DATE: 1997-06-06  
; PRIOR APPLICATION NUMBER: 60/048,894  
; PRIOR FILING DATE: 1997-06-06  
; PRIOR APPLICATION NUMBER: 60/048,971  
; PRIOR FILING DATE: 1997-06-06  
; PRIOR APPLICATION NUMBER: 60/048,964  
; PRIOR FILING DATE: 1997-06-06  
; PRIOR APPLICATION NUMBER: 60/048,882  
; PRIOR FILING DATE: 1997-06-06  
; PRIOR APPLICATION NUMBER: 60/048,899  
; PRIOR FILING DATE: 1997-06-06  
; PRIOR APPLICATION NUMBER: 60/048,893  
; PRIOR FILING DATE: 1997-06-06  
; PRIOR APPLICATION NUMBER: 60/048,900  
; PRIOR FILING DATE: 1997-06-06  
; PRIOR APPLICATION NUMBER: 60/048,901  
; PRIOR FILING DATE: 1997-06-06  
; PRIOR APPLICATION NUMBER: 60/048,892  
; PRIOR FILING DATE: 1997-06-06  
; PRIOR APPLICATION NUMBER: 60/048,915  
; PRIOR FILING DATE: 1997-06-06  
; PRIOR APPLICATION NUMBER: 60/049,019  
; PRIOR FILING DATE: 1997-06-06  
; PRIOR APPLICATION NUMBER: 60/048,970  
; PRIOR FILING DATE: 1997-06-06  
; PRIOR APPLICATION NUMBER: 60/048,972  
; PRIOR FILING DATE: 1997-06-06  
; PRIOR APPLICATION NUMBER: 60/048,916  
; PRIOR FILING DATE: 1997-06-06  
; PRIOR APPLICATION NUMBER: 60/049,373  
; PRIOR FILING DATE: 1997-06-06  
; PRIOR APPLICATION NUMBER: 60/048,949  
; PRIOR FILING DATE: 1997-06-06  
; PRIOR APPLICATION NUMBER: 60/048,974  
; PRIOR FILING DATE: 1997-06-06  
; PRIOR APPLICATION NUMBER: 60/048,883  
; PRIOR FILING DATE: 1997-06-06  
; PRIOR APPLICATION NUMBER: 60/048,897  
; PRIOR FILING DATE: 1997-06-06  
; PRIOR APPLICATION NUMBER: 60/048,898  
; PRIOR FILING DATE: 1997-06-06  
; PRIOR APPLICATION NUMBER: 60/048,962  
; PRIOR FILING DATE: 1997-06-06  
; PRIOR APPLICATION NUMBER: 60/048,963  
; PRIOR FILING DATE: 1997-06-06  
; PRIOR APPLICATION NUMBER: 60/048,877  
; PRIOR FILING DATE: 1997-06-06  
; PRIOR APPLICATION NUMBER: 60/048,878  
; PRIOR FILING DATE: 1997-06-06  
; PRIOR APPLICATION NUMBER: 60/068,054  
; PRIOR FILING DATE: 1997-12-18  
; PRIOR APPLICATION NUMBER: 60/068,064  
; PRIOR FILING DATE: 1997-12-18  
; PRIOR APPLICATION NUMBER: 60/068,053  
; PRIOR FILING DATE: 1997-12-18  
; PRIOR APPLICATION NUMBER: 60/070,923  
; PRIOR FILING DATE: 1997-12-18  
; PRIOR APPLICATION NUMBER: 60/073,160  
; PRIOR FILING DATE: 1998-01-30  
; PRIOR APPLICATION NUMBER: 60/073,159  
; PRIOR FILING DATE: 1998-01-30  
; PRIOR APPLICATION NUMBER: 60/073,165  
; PRIOR FILING DATE: 1998-01-30  
; PRIOR APPLICATION NUMBER: 60/073,164  
; PRIOR FILING DATE: 1998-01-30  
; PRIOR APPLICATION NUMBER: 60/085,925

; PRIOR FILING DATE: 1998-05-18  
; PRIOR APPLICATION NUMBER: 60/085,921  
; PRIOR FILING DATE: 1998-05-18  
; PRIOR APPLICATION NUMBER: 60/085,923  
; PRIOR FILING DATE: 1998-05-18  
; PRIOR APPLICATION NUMBER: 60/085,922  
; PRIOR FILING DATE: 1998-05-18  
; PRIOR APPLICATION NUMBER: 60/092,921  
; PRIOR FILING DATE: 1998-07-15  
; PRIOR APPLICATION NUMBER: 60/094,657  
; PRIOR FILING DATE: 1998-07-30  
; NUMBER OF SEQ ID NOS: 1245  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 1030  
; LENGTH: 66  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-09-933-767-1030  
  
Query Match 59.5%; Score 25; DB 12; Length 66;  
Best Local Similarity 44.4%; Pred. No. 5.9e+02;  
Matches 4; Conservative 3; Mismatches 2; Indels 0; Gaps 0;  
  
Qy 1 NVLMANNII 9  
Db 58 NILKALGIV 66  
  
RESULT 25  
US-10-023-282-1030  
; Sequence 1030, Application US/10023282  
; Publication No. US20030092893A1  
; GENERAL INFORMATION:  
; APPLICANT: Young et al.  
; TITLE OF INVENTION: 207 Human Secreted Proteins  
; FILE REFERENCE: P2007P1  
; CURRENT APPLICATION NUMBER: US/10/023,282  
; CURRENT FILING DATE: 2001-12-20  
; EARLIER APPLICATION NUMBER: 09/205,258  
; EARLIER FILING DATE: 1998-12-04  
; EARLIER APPLICATION NUMBER: PCT/US98/11422  
; EARLIER FILING DATE: 1998-06-04  
; EARLIER APPLICATION NUMBER: 60/048,885  
; EARLIER FILING DATE: 1997-06-06  
; EARLIER APPLICATION NUMBER: 60/049,375  
; EARLIER FILING DATE: 1997-06-06  
; EARLIER APPLICATION NUMBER: 60/048,881  
; EARLIER FILING DATE: 1997-06-06  
; EARLIER APPLICATION NUMBER: 60/048,880  
; EARLIER FILING DATE: 1997-06-06  
; EARLIER APPLICATION NUMBER: 60/048,896  
; EARLIER FILING DATE: 1997-06-06  
; EARLIER APPLICATION NUMBER: 60/049,020  
; EARLIER FILING DATE: 1997-06-06  
; EARLIER APPLICATION NUMBER: 60/048,876  
; EARLIER FILING DATE: 1997-06-06  
; EARLIER APPLICATION NUMBER: 60/048,895  
; EARLIER FILING DATE: 1997-06-06  
; EARLIER APPLICATION NUMBER: 60/048,884  
; EARLIER FILING DATE: 1997-06-06  
; EARLIER APPLICATION NUMBER: 60/048,894  
; EARLIER FILING DATE: 1997-06-06  
; EARLIER APPLICATION NUMBER: 60/048,971  
; EARLIER FILING DATE: 1997-06-06  
; EARLIER APPLICATION NUMBER: 60/048,964  
; EARLIER FILING DATE: 1997-06-06  
; EARLIER APPLICATION NUMBER: 60/048,882  
; EARLIER FILING DATE: 1997-06-06  
; EARLIER APPLICATION NUMBER: 60/048,899  
; EARLIER FILING DATE: 1997-06-06  
; EARLIER APPLICATION NUMBER: 60/048,893  
; EARLIER FILING DATE: 1997-06-06  
; EARLIER APPLICATION NUMBER: 60/048,900

Tue Feb 17 11:55:47 2004

EARLIER FILING DATE: 1997-06-06  
EARLIER APPLICATION NUMBER: 60/048,901  
EARLIER FILING DATE: 1997-06-06  
EARLIER APPLICATION NUMBER: 60/048,892  
EARLIER FILING DATE: 1997-06-06  
EARLIER APPLICATION NUMBER: 60/048,915  
EARLIER FILING DATE: 1997-06-06  
EARLIER APPLICATION NUMBER: 60/049,019  
EARLIER FILING DATE: 1997-06-06  
EARLIER APPLICATION NUMBER: 60/048,970  
EARLIER FILING DATE: 1997-06-06  
EARLIER APPLICATION NUMBER: 60/048,972  
EARLIER FILING DATE: 1997-06-06  
EARLIER APPLICATION NUMBER: 60/048,916  
EARLIER FILING DATE: 1997-06-06  
EARLIER APPLICATION NUMBER: 60/049,373  
EARLIER FILING DATE: 1997-06-06  
EARLIER APPLICATION NUMBER: 60/048,875  
EARLIER FILING DATE: 1997-06-06  
EARLIER APPLICATION NUMBER: 60/049,374  
EARLIER FILING DATE: 1997-06-06  
EARLIER APPLICATION NUMBER: 60/048,883  
EARLIER FILING DATE: 1997-06-06  
EARLIER APPLICATION NUMBER: 60/048,897  
EARLIER FILING DATE: 1997-06-06  
EARLIER APPLICATION NUMBER: 60/048,898  
EARLIER FILING DATE: 1997-06-06  
EARLIER APPLICATION NUMBER: 60/048,962  
EARLIER FILING DATE: 1997-06-06  
EARLIER APPLICATION NUMBER: 60/048,963  
EARLIER FILING DATE: 1997-06-06  
EARLIER APPLICATION NUMBER: 60/048,877  
EARLIER FILING DATE: 1997-06-06  
EARLIER APPLICATION NUMBER: 60/048,878  
EARLIER FILING DATE: 1997-06-06  
EARLIER APPLICATION NUMBER: 60/070,923  
EARLIER FILING DATE: 1997-12-18  
EARLIER APPLICATION NUMBER: 60/092,921  
EARLIER FILING DATE: 1998-07-15  
EARLIER APPLICATION NUMBER: 60/094,657  
EARLIER FILING DATE: 1998-07-30  
NUMBER OF SEQ ID NOS: 1227  
SOFTWARE: PatentIn Ver. 2.0  
SEQ ID NO 1030  
LENGTH: 66  
TYPE: PRT  
ORGANISM: Homo sapiens  
US-10-023-282-1030

Query Match 59.5%; Score 25; DB 15; Length 66;  
Best Local Similarity 44.4%; Pred. No. 5.9e+02;  
Matches 4; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 1 NVLMANNII 9  
DB 58 NILKALGIV 66

RESULT 26  
US-09-864-408A-5276  
Sequence 5276, Application US/09864408A  
Publication No. US2004009474A1  
GENERAL INFORMATION:  
APPLICANT: Leach, Martin D.  
APPLICANT: Shimkets, Richard A.  
TITLE OF INVENTION: No. US2004009474A1 Human Polynucleotides and Polypeptides Encod  
FILE REFERENCE: 21402-012

CURRENT APPLICATION NUMBER: US/09/864,408A  
CURRENT FILING DATE: 2001-05-24  
PRIOR APPLICATION NUMBER: 60/206,690  
PRIOR FILING DATE: 2000-05-24  
NUMBER OF SEQ ID NOS: 9068  
SOFTWARE: FastSeq for Windows Version 4.0  
SEQ ID NO 5276  
LENGTH: 69  
TYPE: PRT  
ORGANISM: Homo sapiens  
US-09-864-408A-5276

Query Match 59.5%; Score 25; DB 12; Length 69;  
Best Local Similarity 71.4%; Pred. No. 6.2e+02;  
Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 NVLMANN 7  
DB 22 NVKMSMN 28

RESULT 27  
US-09-864-761-33623  
Sequence 33623, Application US/09864761  
Patent No. US20020048763A1  
GENERAL INFORMATION:  
APPLICANT: Penn, Sharon G.  
APPLICANT: Rank, David R.  
APPLICANT: Hanzel, David K.  
APPLICANT: Chen, Wensheng  
TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR  
FILE REFERENCE: Aemica-X-1  
CURRENT APPLICATION NUMBER: US/09/864,761  
CURRENT FILING DATE: 2001-05-23  
PRIOR APPLICATION NUMBER: US 60/180,312  
PRIOR FILING DATE: 2000-02-04  
PRIOR APPLICATION NUMBER: US 60/207,456  
PRIOR FILING DATE: 2000-05-26  
PRIOR APPLICATION NUMBER: US 09/632,366  
PRIOR FILING DATE: 2000-08-03  
PRIOR APPLICATION NUMBER: GB 24263.6  
PRIOR FILING DATE: 2000-10-04  
PRIOR APPLICATION NUMBER: US 60/236,359  
PRIOR FILING DATE: 2000-09-27  
PRIOR APPLICATION NUMBER: PCT/US01/00666  
PRIOR FILING DATE: 2001-01-30  
PRIOR APPLICATION NUMBER: PCT/US01/00667  
PRIOR FILING DATE: 2001-01-30  
PRIOR APPLICATION NUMBER: PCT/US01/00664  
PRIOR FILING DATE: 2001-01-30  
PRIOR APPLICATION NUMBER: PCT/US01/00669  
PRIOR FILING DATE: 2001-01-30  
PRIOR APPLICATION NUMBER: PCT/US01/00665  
PRIOR FILING DATE: 2001-01-30  
PRIOR APPLICATION NUMBER: PCT/US01/00668  
PRIOR FILING DATE: 2001-01-30  
PRIOR APPLICATION NUMBER: PCT/US01/00663  
PRIOR FILING DATE: 2001-01-30  
PRIOR APPLICATION NUMBER: PCT/US01/00662  
PRIOR FILING DATE: 2001-01-30  
PRIOR APPLICATION NUMBER: PCT/US01/00661  
PRIOR FILING DATE: 2001-01-30  
PRIOR APPLICATION NUMBER: PCT/US01/00670  
PRIOR FILING DATE: 2001-01-30  
PRIOR APPLICATION NUMBER: US 60/234,687  
PRIOR FILING DATE: 2000-09-21  
PRIOR APPLICATION NUMBER: US 09/608,408  
PRIOR FILING DATE: 2000-06-30  
PRIOR APPLICATION NUMBER: US 09/774,203  
PRIOR FILING DATE: 2001-01-29  
NUMBER OF SEQ ID NOS: 49117  
SOFTWARE: Annomax Sequence Listing Engine vers. 1.1



```
; SEQ ID NO 33623
; LENGTH: 78
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: MAP TO AP00201.1
; OTHER INFORMATION: EXPRESSED IN HEL100, SIGNAL = 1
; OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 1.2
; OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 0.83
; OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 1.8
; OTHER INFORMATION: EXPRESSED IN BT474, SIGNAL = 1.3
; OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 0.96
; OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 1.1
; OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 1.1
; OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 26
; OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 1
; OTHER INFORMATION: EST HUMAN HIT: AA402241.1, EVALUATE 9.00e-34
; OTHER INFORMATION: SWISSPROT HIT: O14867, EVALUATE 3.00e-39
US-09-864-761-33623

Query Match 59.5%; Score 25; DB 9; Length 78;
Best Local Similarity 57.1%; Pred. No. 7.1e+02;
Matches 4; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

Qy 1 NVLMAMN 7
Db 19 NVLLSLN 25

RESULT 28
US-09-933-767-1032
; Sequence 1032, Application US/099333767
; Publication No. US20030181692A1
; GENERAL INFORMATION:
; APPLICANT: Ni et al.
; TITLE OF INVENTION: 207 Human Secreted Proteins
; FILE REFERENCE: P2007P2
; CURRENT APPLICATION NUMBER: US/09/933.767
; CURRENT FILING DATE: 2001-08-22
; PCT APPLICATION NUMBER: PCT/US01/05614
; PRIOR FILING DATE: 2001-02-21
; PRIOR APPLICATION NUMBER: 60/184,836
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: 60/193,170
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: 09/205,258
; PRIOR FILING DATE: 1998-12-04
; PRIOR APPLICATION NUMBER: PCT/US98/11422
; PRIOR FILING DATE: 1998-06-04
; PRIOR APPLICATION NUMBER: 60/048,885
; PRIOR FILING DATE: 1997-06-06
; PRIOR APPLICATION NUMBER: 60/049,375
; PRIOR FILING DATE: 1997-06-06
; PRIOR APPLICATION NUMBER: 60/048,881
; PRIOR FILING DATE: 1997-06-06
; PRIOR APPLICATION NUMBER: 60/048,880
; PRIOR FILING DATE: 1997-06-06
; PRIOR APPLICATION NUMBER: 60/048,896
; PRIOR FILING DATE: 1997-06-06
; PRIOR APPLICATION NUMBER: 60/049,020
; PRIOR FILING DATE: 1997-06-06
; PRIOR APPLICATION NUMBER: 60/048,876
; PRIOR FILING DATE: 1997-06-06
; PRIOR APPLICATION NUMBER: 60/048,895
; PRIOR FILING DATE: 1997-06-06
; PRIOR APPLICATION NUMBER: 60/048,884
; PRIOR FILING DATE: 1997-06-06
; PRIOR APPLICATION NUMBER: 60/048,894
; PRIOR FILING DATE: 1997-06-06
; PRIOR APPLICATION NUMBER: 60/048,971
; PRIOR FILING DATE: 1997-06-06
; PRIOR APPLICATION NUMBER: 60/048,964
; PRIOR FILING DATE: 1997-06-06
; PRIOR APPLICATION NUMBER: 60/048,882
; PRIOR FILING DATE: 1997-06-06
; PRIOR APPLICATION NUMBER: 60/048,899
; PRIOR FILING DATE: 1997-06-06
; PRIOR APPLICATION NUMBER: 60/048,893
; PRIOR FILING DATE: 1997-06-06
; PRIOR APPLICATION NUMBER: 60/048,900
; PRIOR FILING DATE: 1997-06-06
; PRIOR APPLICATION NUMBER: 60/048,901
; PRIOR FILING DATE: 1997-06-06
; PRIOR APPLICATION NUMBER: 60/048,892
; PRIOR FILING DATE: 1997-06-06
; PRIOR APPLICATION NUMBER: 60/048,915
; PRIOR FILING DATE: 1997-06-06
; PRIOR APPLICATION NUMBER: 60/049,019
; PRIOR FILING DATE: 1997-06-06
; PRIOR APPLICATION NUMBER: 60/048,970
; PRIOR FILING DATE: 1997-06-06
; PRIOR APPLICATION NUMBER: 60/048,972
; PRIOR FILING DATE: 1997-06-06
; PRIOR APPLICATION NUMBER: 60/048,916
; PRIOR FILING DATE: 1997-06-06
; PRIOR APPLICATION NUMBER: 60/049,373
; PRIOR FILING DATE: 1997-06-06
; PRIOR APPLICATION NUMBER: 60/048,875
; PRIOR FILING DATE: 1997-06-06
; PRIOR APPLICATION NUMBER: 60/049,374
; PRIOR FILING DATE: 1997-06-06
; PRIOR APPLICATION NUMBER: 60/048,974
; PRIOR FILING DATE: 1997-06-06
; PRIOR APPLICATION NUMBER: 60/048,883
; PRIOR FILING DATE: 1997-06-06
; PRIOR APPLICATION NUMBER: 60/048,897
; PRIOR FILING DATE: 1997-06-06
; PRIOR APPLICATION NUMBER: 60/048,898
; PRIOR FILING DATE: 1997-06-06
; PRIOR APPLICATION NUMBER: 60/048,962
; PRIOR FILING DATE: 1997-06-06
; PRIOR APPLICATION NUMBER: 60/048,963
; PRIOR FILING DATE: 1997-06-06
; PRIOR APPLICATION NUMBER: 60/048,877
; PRIOR FILING DATE: 1997-06-06
; PRIOR APPLICATION NUMBER: 60/048,878
; PRIOR FILING DATE: 1997-06-06
; PRIOR APPLICATION NUMBER: 60/068,054
; PRIOR FILING DATE: 1997-12-18
; PRIOR APPLICATION NUMBER: 60/068,064
; PRIOR FILING DATE: 1997-12-18
; PRIOR APPLICATION NUMBER: 60/068,053
; PRIOR FILING DATE: 1997-12-18
; PRIOR APPLICATION NUMBER: 60/070,923
; PRIOR FILING DATE: 1997-12-18
; PRIOR APPLICATION NUMBER: 60/073,160
; PRIOR FILING DATE: 1998-01-30
; PRIOR APPLICATION NUMBER: 60/073,159
; PRIOR FILING DATE: 1998-01-30
; PRIOR APPLICATION NUMBER: 60/073,165
; PRIOR FILING DATE: 1998-01-30
; PRIOR APPLICATION NUMBER: 60/073,164
; PRIOR FILING DATE: 1998-01-30
; PRIOR APPLICATION NUMBER: 60/085,925
; PRIOR FILING DATE: 1998-05-18
; PRIOR APPLICATION NUMBER: 60/085,921
; PRIOR FILING DATE: 1998-05-18
; PRIOR APPLICATION NUMBER: 60/085,923
; PRIOR FILING DATE: 1998-05-18
; PRIOR APPLICATION NUMBER: 60/085,922
; PRIOR FILING DATE: 1998-05-18
; PRIOR APPLICATION NUMBER: 60/092,921
```

; PRIOR FILING DATE: 1998-07-15
; PRIOR APPLICATION NUMBER: 60/094,657
; PRIOR FILING DATE: 1998-07-30
; NUMBER OF SEQ ID NOS: 1245
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1032
; LENGTH: 84
; TYPE: PRT
; ORGANISM: Homo sapiens
; US-09-933-767-1032

Query Match 59.5%; Score 25; DB 12; Length 84;
Best Local Similarity 44.4%; Pred. No. 7.7e+02;
Matches 4; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

Qy 1 NVLMAMNII 9
|:| |:|
Db 71 NILKALGIV 79

RESULT 29
US-10-023-282-1032
; Sequence 1032, Application US/10023282
; Publication No. US20030092893A1
; GENERAL INFORMATION:
; APPLICANT: Young et al.
; TITLE OF INVENTION: 207 Human Secreted Proteins
; FILE REFERENCE: P2007P1
; CURRENT APPLICATION NUMBER: US/10/023,282
; CURRENT FILING DATE: 2001-12-20
; EARLIER APPLICATION NUMBER: 09/205,258
; EARLIER FILING DATE: 1998-12-04
; EARLIER APPLICATION NUMBER: PCT/US98/11422
; EARLIER FILING DATE: 1998-06-04
; EARLIER APPLICATION NUMBER: 60/048,885
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/049,375
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,881
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,880
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,896
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/049,020
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,876
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,895
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,884
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,894
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,971
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,964
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,882
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,899
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,893
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,900
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,901
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,892
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,915
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/049,019

; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,970
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,972
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,916
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/049,373
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,875
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/049,374
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,917
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,949
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,974
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,883
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,897
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,898
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,962
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,963
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,877
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,878
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/070,923
; EARLIER FILING DATE: 1997-12-18
; EARLIER APPLICATION NUMBER: 60/092,921
; EARLIER FILING DATE: 1998-07-15
; EARLIER APPLICATION NUMBER: 60/094,657
; EARLIER FILING DATE: 1998-07-30
; NUMBER OF SEQ ID NOS: 1227
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1032
; LENGTH: 84
; TYPE: PRT
; ORGANISM: Homo sapiens
; US-10-023-282-1032

Query Match 59.5%; Score 25; DB 15; Length 84;
Best Local Similarity 44.4%; Pred. No. 7.7e+02;
Matches 4; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

Qy 1 NVLMAMNII 9
|:| |:|
Db 71 NILKALGIV 79

RESULT 30
US-10-029-386-33793
; Sequence 33793, Application US/10029386
; Publication No. US20030194704A1
; GENERAL INFORMATION:
; APPLICANT: Penn, Sharron G.
; APPLICANT: Hanzel, David K.
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
; FILE REFERENCE: AEOMICA-X-2
; CURRENT APPLICATION NUMBER: US/10/029,386
; CURRENT FILING DATE: 2001-12-20
; NUMBER OF SEQ ID NOS: 34288
; SOFTWARE: Annonmax Sequence Listing Engine vers. 1.1
; SEQ ID NO 33793
; LENGTH: 88

```
/ TYPE: PRT
/ ORGANISM: Homo sapiens
/ FEATURE:
/ OTHER INFORMATION: MAP TO AC010522.3
/ OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 0.8
/ OTHER INFORMATION: SWISSPROT HIT: O43296, EVALUE 9.90e-02
US-10-029-386-33793

Query Match      59.5%; Score 25; DB 12; Length 88;
Best Local Similarity 50.0%; Pred. No. 8e+02;
Matches 4; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY      1 NVLMANNI 8
Db      20 NVNCAVNV 27

RESULT 31
US-09-864-761-44843
/ Sequence 44843, Application US/09864761
/ Patent No. US20020048763A1
/ GENERAL INFORMATION:
/ APPLICANT: Penn, Sharron G.
/ APPLICANT: Rank, David R.
/ APPLICANT: Hanzel, David K.
/ APPLICANT: Chen, Wensheng
/ TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
/ FILE REFERENCE: Aeomica-X-1
/ CURRENT APPLICATION NUMBER: US/09/864,761
/ CURRENT FILING DATE: 2001-05-23
/ PRIOR APPLICATION NUMBER: US 60/180,312
/ PRIOR FILING DATE: 2000-02-04
/ PRIOR APPLICATION NUMBER: US 60/207,456
/ PRIOR FILING DATE: 2000-05-26
/ PRIOR APPLICATION NUMBER: US 09/632,366
/ PRIOR FILING DATE: 2000-08-03
/ PRIOR APPLICATION NUMBER: GB 24263.6
/ PRIOR FILING DATE: 2000-10-04
/ PRIOR APPLICATION NUMBER: US 60/236,359
/ PRIOR FILING DATE: 2000-09-27
/ PRIOR APPLICATION NUMBER: PCT/US01/00666
/ PRIOR FILING DATE: 2001-01-30
/ PRIOR APPLICATION NUMBER: PCT/US01/00667
/ PRIOR FILING DATE: 2001-01-30
/ PRIOR APPLICATION NUMBER: PCT/US01/00664
/ PRIOR FILING DATE: 2001-01-30
/ PRIOR APPLICATION NUMBER: PCT/US01/00669
/ PRIOR FILING DATE: 2001-01-30
/ PRIOR APPLICATION NUMBER: PCT/US01/00665
/ PRIOR FILING DATE: 2001-01-30
/ PRIOR APPLICATION NUMBER: PCT/US01/00668
/ PRIOR FILING DATE: 2001-01-30
/ PRIOR APPLICATION NUMBER: PCT/US01/00663
/ PRIOR FILING DATE: 2001-01-30
/ PRIOR APPLICATION NUMBER: PCT/US01/00662
/ PRIOR FILING DATE: 2001-01-30
/ PRIOR APPLICATION NUMBER: PCT/US01/00661
/ PRIOR FILING DATE: 2001-01-30
/ PRIOR APPLICATION NUMBER: PCT/US01/00670
/ PRIOR FILING DATE: 2001-01-30
/ PRIOR APPLICATION NUMBER: US 60/234,687
/ PRIOR FILING DATE: 2000-09-21
/ PRIOR APPLICATION NUMBER: US 09/608,408
/ PRIOR FILING DATE: 2000-06-30
/ PRIOR APPLICATION NUMBER: US 09/774,203
/ PRIOR FILING DATE: 2001-01-29
/ NUMBER OF SEQ ID NOS: 49117
/ SOFTWARE: Annonax Sequence Listing Engine vers. 1.1
/ SEQ ID NO 44843
/ LENGTH: 64
/ TYPE: PRT
/ ORGANISM: Homo sapiens
```

```
/ FEATURE:
/ OTHER INFORMATION: MAP TO AC010458.2
/ OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 0.93
/ OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 0.76
/ OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 0.7
/ OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 1.1
/ OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 1.3
/ OTHER INFORMATION: EST HUMAN HIT: A1650936.1, EVALUE 6.00e-29
/ OTHER INFORMATION: SWISSPROT HIT: P49454, EVALUE 5.90e+00
US-09-864-761-44843

Query Match      58.3%; Score 24.5; DB 9; Length 64;
Best Local Similarity 70.0%; Pred. No. 7.2e+02;
Matches 7; Conservative 1; Mismatches 1; Indels 1; Gaps 1;

QY      1 NVLMA-MNII 9
Db      39 NVQARMNVI 48

RESULT 32
US-10-376-121A-66
/ Sequence 66, Application US/10376121A
/ Publication No. US20030216544A1
/ GENERAL INFORMATION:
/ APPLICANT: Hartley, John
/ TITLE OF INVENTION: METHODS AND REAGENTS FOR DIAGNOSIS OF
/ AUTOANTIBODIES
/ NUMBER OF SEQUENCES: 218
/ CORRESPONDENCE ADDRESS:
/ ADDRESSEE: Patrea L. Pabst
/ STREET: Suite 2000, 1201 West Peachtree Street, N.E.
/ CITY: Atlanta
/ STATE: GA
/ COUNTRY: USA
/ ZIP: 30309-3400
/ COMPUTER READABLE FORM:
/ MEDIUM TYPE: Floppy disk
/ COMPUTER: IBM PC compatible
/ OPERATING SYSTEM: PC-DOS/MS-DOS
/ SOFTWARE: PatentIn Release #1.0, Version #1.25
/ CURRENT APPLICATION DATA:
/ APPLICATION NUMBER: US/10/376,121A
/ FILING DATE: 27-Mar-2003
/ CLASSIFICATION: <Unknown>
/ PRIOR APPLICATION DATA:
/ APPLICATION NUMBER: 07/867,819
/ FILING DATE: April 13, 1992
/ APPLICATION NUMBER: 07/648,205
/ FILING DATE: January 31, 1991
/ APPLICATION NUMBER: 07/472,947
/ FILING DATE: January 31, 1990
/ ATTORNEY/AGENT INFORMATION:
/ NAME: Pabst, Patrea L.
/ REGISTRATION NUMBER: 31,284
/ REFERENCE/DOCKET NUMBER: OMRF114CIP(2)DIV(2)
/ TELECOMMUNICATION INFORMATION:
/ TELEPHONE: (404)-817-8473
/ TELEFAX: (404)-817-8588
/ INFORMATION FOR SEQ ID NO: 66:
/ SEQUENCE CHARACTERISTICS:
/ LENGTH: 8 amino acids
/ TYPE: amino acid
/ STRANDEDNESS: single
/ TOPOLOGY: linear
/ MOLECULE TYPE: peptide
/ FEATURE:
/ NAME/KEY: Binding-site
/ LOCATION: 1..8
/ SEQUENCE DESCRIPTION: SEQ ID NO: 66:
US-10-376-121A-66

Query Match      57.1%; Score 24; DB 12; Length 8;
```

**Tue Feb 17 11:55:47 2004**

**us-09-900-147-2.rapb**

Best Local Similarity	75.0%;	Pred. No. 7e+05;	
Matches	6;	Conservative	1; Indels
		Mismatches	0; Gaps

Qy 2 VLMAMNII 9  
{}{}{}{}: {}  
pb 1 VLMAMSOI 8

RESULT 33

```

US-09-864-761-46815
; Sequence 46815, Application US/09864761
; Patent No. US20020048763A1
; GENERAL INFORMATION:
; APPLICANT: Penn, Sharron G.
; APPLICANT: Rank, David R.
; APPLICANT: Hanzel, David K.
; APPLICANT: Chen, Wensheng
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
; FILE REFERENCE: Aeomica-X-1
; CURRENT APPLICATION NUMBER: US/09/864,761
; CURRENT FILING DATE: 2001-05-23
; PRIOR APPLICATION NUMBER: US 60/180,312
; PRIOR FILING DATE: 2000-02-04
; PRIOR APPLICATION NUMBER: US 60/207,456
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: US 09/632,366
; PRIOR FILING DATE: 2000-08-03
; PRIOR APPLICATION NUMBER: GB 24263.6
; PRIOR FILING DATE: 2000-10-04
; PRIOR APPLICATION NUMBER: US 60/236,359
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: PCT/US01/00666
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00667
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00664
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00669
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00665
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00668
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00663
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00662
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00661
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00670
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: US 60/234,687
; PRIOR FILING DATE: 2000-09-21
; PRIOR APPLICATION NUMBER: US 09/608,408
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: US 09/774,203
; PRIOR FILING DATE: 2001-01-29
; NUMBER OF SEQ ID NOS: 49117
; SOFTWARE: Annomax Sequence Listing Engine vers. 1.1
; SEQ ID NO 46815
; LENGTH: 25
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: MAP TO AC007130.2
; OTHER INFORMATION: EXPRESSED IN BT474, SIGNAL = 0.74
; OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 0.8
; OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 0.58
; OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 0.59
; OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 0.84
; OTHER INFORMATION: EST HUMAN HIT: BF340337.1, EVALUE 3.00e-06
; OTHER INFORMATION: SWISSPROT HIT: P29266, EVALUE 5.00e-07

```

US-09-864-761-46815

```
Query Match          57.1%; Score 24; DB 9; Length 25;
Best Local Similarity 44.4%; Pred. No. 3.3e+02;
Matches 4; Conservative 5; Mismatches 0; Indels 0; Gaps 0;
```

Qy 1 NVLMAMNII 9  
|:|:|:|:|  
Db 7 NMLLAISM I 15

RESULT 34

RE3302. 1  
 US-10-120-604-112  
 ; Sequence 112, Application US/10120604  
 ; Publication No. US20030096347A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Bristol-Myers Squibb Company  
 ; TITLE OF INVENTION: POLYNUCLEOTIDES ENCODING TWO NOVEL HUMAN G-PROTEIN COUPLED RECEPTOR  
 ; TITLE OF INVENTION: HGPBMY28 AND HGPBMY29, AND SPLICE VARIANTS THEREOF  
 ; FILE REFERENCE: D0143NP  
 ; CURRENT APPLICATION NUMBER: US/10/120,604  
 ; CURRENT FILING DATE: 2002-04-11  
 ; PRIOR APPLICATION NUMBER: US 60/283,145  
 ; PRIOR FILING DATE: 2001-04-11  
 ; PRIOR APPLICATION NUMBER: US 60/283,161  
 ; PRIOR FILING DATE: 2001-04-11  
 ; PRIOR APPLICATION NUMBER: US 60/288,468  
 ; PRIOR FILING DATE: 2001-05-03  
 ; PRIOR APPLICATION NUMBER: US 60/300,619  
 ; PRIOR FILING DATE: 2001-06-25  
 ; NUMBER OF SEQ ID NOS: 226  
 ; SOFTWARE: PatentIn version 3.0  
 ; SEQ ID NO 112  
 ; LENGTH: 25  
 ; TYPE: PRT  
 ; ORGANISM: Homo sapiens  
 ; US-10-120-604-112

Query Match	57.1%	Score 24;	DB 15;	Length 25;	.
Best Local Similarity	50.0%;	Pred. NO. 3.3e+02;			
Matches	4;	Conservative	4;	Mismatches	0;
				Indels	0;
				Gaps	0;

Qy 2 VLMAMNII 9  
|::|::|::|  
nb 11 VIVAVNLI 18

RESULT 35

US-09-864-761-40663  
; Sequence 40663, Application US/09864761  
; Patent No. US20020048763A1  
; GENERAL INFORMATION:  
; APPLICANT: Penn, Sharron G.  
; APPLICANT: Rank, David K.  
; APPLICANT: Hanzel, David K.  
; APPLICANT: Chen, Wensheng  
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR  
; TITLE OF INVENTION: GENE EXPRESSION ANALYSIS BY MICROARRAY  
; FILE REFERENCE: Aecmica-X-1  
; CURRENT APPLICATION NUMBER: US/09/864,761  
; CURRENT FILING DATE: 2001-05-23  
; PRIOR APPLICATION NUMBER: US 60/180,312  
; PRIOR FILING DATE: 2000-02-04  
; PRIOR APPLICATION NUMBER: US 60/207,456  
; PRIOR FILING DATE: 2000-05-26  
; PRIOR APPLICATION NUMBER: US 09/632,366  
; PRIOR FILING DATE: 2000-08-03  
; PRIOR APPLICATION NUMBER: GB 24263.6  
; PRIOR FILING DATE: 2000-10-04  
; PRIOR APPLICATION NUMBER: US 60/236,359  
; PRIOR FILING DATE: 2000-09-27  
; PRIOR APPLICATION NUMBER: PCT/US01/00666  
; PRIOR FILING DATE: 2001-01-30

;; PRIOR APPLICATION NUMBER: PCT/US01/00667  
;; PRIOR FILING DATE: 2001-01-30  
;; PRIOR APPLICATION NUMBER: PCT/US01/00664  
;; PRIOR FILING DATE: 2001-01-30  
;; PRIOR APPLICATION NUMBER: PCT/US01/00669  
;; PRIOR FILING DATE: 2001-01-30  
;; PRIOR APPLICATION NUMBER: PCT/US01/00665  
;; PRIOR FILING DATE: 2001-01-30  
;; PRIOR APPLICATION NUMBER: PCT/US01/00668  
;; PRIOR FILING DATE: 2001-01-30  
;; PRIOR APPLICATION NUMBER: PCT/US01/00663  
;; PRIOR FILING DATE: 2001-01-30  
;; PRIOR APPLICATION NUMBER: PCT/US01/00662  
;; PRIOR FILING DATE: 2001-01-30  
;; PRIOR APPLICATION NUMBER: PCT/US01/00661  
;; PRIOR FILING DATE: 2001-01-30  
;; PRIOR APPLICATION NUMBER: PCT/US01/00670  
;; PRIOR FILING DATE: 2001-01-30  
;; PRIOR APPLICATION NUMBER: US 60/234,687  
;; PRIOR FILING DATE: 2000-09-21  
;; PRIOR APPLICATION NUMBER: US 09/608,408  
;; PRIOR FILING DATE: 2000-06-30  
;; PRIOR APPLICATION NUMBER: US 09/774,203  
;; PRIOR FILING DATE: 2001-01-29  
;; NUMBER OF SEQ ID NOS: 49117  
;; SOFTWARE: Annomax Sequence Listing Engine vers. 1.1  
;; SEQ ID NO 40663  
;; LENGTH: 28  
;; TYPE: PRT  
;; ORGANISM: Homo sapiens  
;; FEATURE:  
;; OTHER INFORMATION: MAP TO AL133319.21  
;; OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 1.9  
;; OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 2.3  
;; OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 4  
;; OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 2.5  
;; OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 3.3  
;; OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 2.3  
;; OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 2.1  
;; OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 3.2  
US-09-864-761-40663

Query Match 57.1%; Score 24; DB 9; Length 28;  
Best Local Similarity 55.6%; Pred. No. 3.8e+02;  
Matches 5; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

Qy 1 NVLMANNII 9  
Db 5 SIKDMNII 13

RESULT 36  
US-10-387-977-27  
;; Sequence 27, Application US/10387977  
;; Publication No. US20040005276A1  
;; GENERAL INFORMATION:  
;; APPLICANT: Reynolds, Eric Charles  
;; APPLICANT: O'Brien-Simpson, Neil Martin  
;; APPLICANT: Skakeski, Nada  
;; TITLE OF INVENTION: SYNTHETIC PEPTIDE CONSTRUCTS FOR THE  
;; TITLE OF INVENTION: DIAGNOSIS AND TREATMENT OF PERIODONTITIS ASSOCIATED WITH  
;; TITLE OF INVENTION: PORPHYROMONAS GINGIVALIS  
;; FILE REFERENCE: 529282000301  
;; CURRENT APPLICATION NUMBER: US/10/387,977  
;; CURRENT FILING DATE: 2003-07-18  
;; PRIOR APPLICATION NUMBER: US 09/423,056  
;; PRIOR FILING DATE: 2000-03-22  
;; PRIOR APPLICATION NUMBER: PCT/AU98/00311  
;; PRIOR FILING DATE: 1998-04-30  
;; PRIOR APPLICATION NUMBER: AU PO 6528  
;; PRIOR FILING DATE: 1997-04-30  
;; NUMBER OF SEQ ID NOS: 105  
;; SOFTWARE: FastSeq for Windows Version 4.0

;; SEQ ID NO 27  
;; LENGTH: 29  
;; TYPE: PRT  
;; ORGANISM: Porphyromonas gingivalis  
US-10-387-977-27  
Query Match 57.1%; Score 24; DB 12; Length 29;  
Best Local Similarity 50.0%; Pred. No. 3.9e+02;  
Matches 4; Conservative 3; Mismatches 1; Indels 0; Gaps 0;  
Qy 1 NVLMANNII 8  
Db 10 NVVIAQNL 17  
RESULT 37  
US-10-224-260-6  
;; Sequence 6, Application US/10224260  
;; Publication No. US20030059845A1  
;; GENERAL INFORMATION:  
;; APPLICANT: Van Tol, Hubert H.M.  
;; APPLICANT: Civelli, Olivier  
;; TITLE OF INVENTION: A No. US20030059845A1el Human Dopamine Receptor and Uses  
;; NUMBER OF SEQUENCES: 22  
;; CORRESPONDENCE ADDRESS:  
;; ADDRESSEE: Allegretti & Witcoff, Ltd.  
;; STREET: 10 South Wacker Drive, Suite 3000  
;; CITY: Chicago  
;; STATE: Illinois  
;; COUNTRY: USA  
;; ZIP: 60606  
;; COMPUTER READABLE FORM:  
;; MEDIUM TYPE: Floppy disk  
;; COMPUTER: IBM PC compatible  
;; OPERATING SYSTEM: PC-DOS/MS-DOS  
;; SOFTWARE: PatentIn Release #1.0, Version #1.25  
;; CURRENT APPLICATION DATA:  
;; APPLICATION NUMBER: US/10/224,260  
;; FILING DATE: 20-Aug-2002  
;; CLASSIFICATION: <Unknown>  
;; PRIOR APPLICATION DATA:  
;; APPLICATION NUMBER: 07/928,611  
;; FILING DATE: <Unknown>  
;; ATTORNEY/AGENT INFORMATION:  
;; NAME: No. US20030059845A1han, Kevin E  
;; REGISTRATION NUMBER: 35,303  
;; REFERENCE/DOCKET NUMBER: 90,1092-B  
;; TELECOMMUNICATION INFORMATION:  
;; TELEPHONE: 312-715-1000  
;; TELEFAX: 312-715-1234  
;; TELEX: 810-221-8317  
;; INFORMATION FOR SEQ ID NO: 6:  
;; SEQUENCE CHARACTERISTICS:  
;; LENGTH: 37 amino acids  
;; TYPE: amino acid  
;; TOPOLOGY: linear  
;; MOLECULE TYPE: protein  
;; SEQUENCE DESCRIPTION: SEQ ID NO: 6:  
US-10-224-260-6

Query Match 57.1%; Score 24; DB 15; Length 37;  
Best Local Similarity 44.4%; Pred. No. 5e+02;  
Matches 4; Conservative 4; Mismatches 1; Indels 0; Gaps 0;

Qy 1 NVLMANNII 9  
Db 14 DALMAMDVM 22

RESULT 38  
US-09-864-761-40202  
;; Sequence 40202, Application US/09864761  
;; Patent No. US20020048763A1

GENERAL INFORMATION:  
; APPLICANT: Penn, Sharon G.  
; APPLICANT: Rank, David R.  
; APPLICANT: Hanzel, David K.  
; APPLICANT: Chen, Wensheng  
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR  
; FILE REFERENCE: Aeonica-X-1  
; CURRENT APPLICATION NUMBER: US 09/864,761  
; CURRENT FILING DATE: 2001-05-23  
; PRIOR APPLICATION NUMBER: US 60/180,312  
; PRIOR FILING DATE: 2000-02-04  
; PRIOR APPLICATION NUMBER: US 60/207,456  
; PRIOR FILING DATE: 2000-05-26  
; PRIOR APPLICATION NUMBER: US 09/632,366  
; PRIOR FILING DATE: 2000-08-03  
; PRIOR APPLICATION NUMBER: GB 24263.6  
; PRIOR FILING DATE: 2000-10-04  
; PRIOR APPLICATION NUMBER: US 60/236,359  
; PRIOR FILING DATE: 2000-09-27  
; PRIOR APPLICATION NUMBER: PCT/US01/00666  
; PRIOR FILING DATE: 2001-01-30  
; PRIOR APPLICATION NUMBER: PCT/US01/00667  
; PRIOR FILING DATE: 2001-01-30  
; PRIOR APPLICATION NUMBER: PCT/US01/00664  
; PRIOR FILING DATE: 2001-01-30  
; PRIOR APPLICATION NUMBER: PCT/US01/00669  
; PRIOR FILING DATE: 2001-01-30  
; PRIOR APPLICATION NUMBER: PCT/US01/00665  
; PRIOR FILING DATE: 2001-01-30  
; PRIOR APPLICATION NUMBER: PCT/US01/00668  
; PRIOR FILING DATE: 2001-01-30  
; PRIOR APPLICATION NUMBER: PCT/US01/00663  
; PRIOR FILING DATE: 2001-01-30  
; PRIOR APPLICATION NUMBER: PCT/US01/00662  
; PRIOR FILING DATE: 2001-01-30  
; PRIOR APPLICATION NUMBER: PCT/US01/00661  
; PRIOR FILING DATE: 2001-01-30  
; PRIOR APPLICATION NUMBER: PCT/US01/00670  
; PRIOR FILING DATE: 2001-01-30  
; PRIOR APPLICATION NUMBER: US 60/234,687  
; PRIOR FILING DATE: 2000-09-21  
; PRIOR APPLICATION NUMBER: US 09/608,408  
; PRIOR FILING DATE: 2000-06-30  
; PRIOR APPLICATION NUMBER: US 09/774,203  
; PRIOR FILING DATE: 2001-01-29  
; NUMBER OF SEQ ID NOS: 49117  
; SOFTWARE: Anomax Sequence Listing Engine vers. 1.1  
; SEQ ID NO 40202  
; LENGTH: 41  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
; FEATURE:  
; OTHER INFORMATION: MAP TO AC021000.3  
; OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 1.5  
; OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 1.4  
; OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 1.7  
; OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 5.2  
; OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 1.8  
; OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 1.8  
; OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 2.6  
; OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 1.8  
; OTHER INFORMATION: EST HUMAN HIT: AV724222.1, EVALUATE 9.00e-17  
; OTHER INFORMATION: SWISSPROT HIT: P54938, EVALUATE 7.00e-18  
US-09-864-761-40202

Query Match 57.1%; Score 24; DB 9; Length 41;  
Best Local Similarity 62.5%; Pred. No. 5.6e+02;  
Matches 5; Conservative 2; Mismatches 1; Indels 0; Gaps 0;  
QY 1 NVLMAMNI 8  
DB 29 NVKMALDI 36

RESULT 39  
US-10-321-857-54  
; Sequence 54, Application US/10321857  
; Publication No. US20030180816A1  
; GENERAL INFORMATION:  
; APPLICANT: Applied NanoSystems B.V.  
; TITLE OF INVENTION: A method to provide bacterial ghosts with antigens  
; FILE REFERENCE: 2183-5547US  
; CURRENT APPLICATION NUMBER: US/10/321,857  
; CURRENT FILING DATE: 2002-12-16  
; PRIOR APPLICATION NUMBER: PCT/NL02/00383  
; PRIOR FILING DATE: 2002-06-11  
; PRIOR APPLICATION NUMBER: 01202239.8  
; PRIOR FILING DATE: 2001-06-11  
; NUMBER OF SEQ ID NOS: 165  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 54  
; LENGTH: 44  
; TYPE: PRT  
; ORGANISM: Haemophilus influenzae  
; FEATURE:  
; NAME/KEY: SITE  
; LOCATION: (1)..(44)  
; OTHER INFORMATION: Acma cell wall binding domain homologue  
US-10-321-857-54  
Query Match 57.1%; Score 24; DB 12; Length 44;  
Best Local Similarity 42.9%; Pred. No. 6.1e+02;  
Matches 3; Conservative 4; Mismatches 0; Indels 0; Gaps 0;  
QY 1 NVLMAMN 7  
DB 21 NILLSLN 27

RESULT 40  
US-10-318-675-54  
; Sequence 54, Application US/10318675  
; Publication No. US20030186851A1  
; GENERAL INFORMATION:  
; APPLICANT: Applied NanoSystems B.V.  
; TITLE OF INVENTION: Bacterial ghosts provided with antigens  
; FILE REFERENCE: 2183-5546US  
; CURRENT APPLICATION NUMBER: US/10/318,675  
; CURRENT FILING DATE: 2002-12-13  
; PRIOR APPLICATION NUMBER: PCT/NL02/00383  
; PRIOR FILING DATE: 2002-06-11  
; PRIOR APPLICATION NUMBER: 01202239.8  
; PRIOR FILING DATE: 2001-06-11  
; NUMBER OF SEQ ID NOS: 165  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 54  
; LENGTH: 44  
; TYPE: PRT  
; ORGANISM: Haemophilus influenzae  
; FEATURE:  
; NAME/KEY: SITE  
; LOCATION: (1)..(44)  
; OTHER INFORMATION: Acma cell wall binding domain homologue  
US-10-318-675-54  
Query Match 57.1%; Score 24; DB 12; Length 44;  
Best Local Similarity 42.9%; Pred. No. 6.1e+02;  
Matches 3; Conservative 4; Mismatches 0; Indels 0; Gaps 0;  
QY 1 NVLMAMN 7  
DB 21 NILLSLN 27

US-10-231-417-202  
; Sequence 202, Application US/10231417  
; Publication No. US20030176681A1  
; GENERAL INFORMATION:  
; APPLICANT: Feng et al.  
; TITLE OF INVENTION: 148 Human Secreted Proteins  
; FILE REFERENCE: P2019P1  
; CURRENT APPLICATION NUMBER: US/10/231.417  
; CURRENT FILING DATE: 2002-08-30  
; PRIOR APPLICATION NUMBER: US/09/296,622  
; PRIOR FILING DATE: 1999-04-23  
; NUMBER OF SEQ ID NOS: 619  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 202  
; LENGTH: 52  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: SITE  
; LOCATION: (52)  
; OTHER INFORMATION: Xaa equals stop translation  
US-10-231-417-202

Query Match 57.1%; Score 24; DB 12; Length 52;  
Best Local Similarity 33.3%; Pred. No. 7.3e+02;  
Matches 3; Conservative 6; Mismatches 0; Indels 0; Gaps 0;

Qy 1 NVLMAMNII 9  
:|:|:|:  
Db 8 SLIVLVNL 16

## RESULT 42

US-09-864-761-36764  
; Sequence 36764, Application US/09864761  
; Patent No. US20020048763A1  
; GENERAL INFORMATION:  
; APPLICANT: Penn, Sharon G.  
; APPLICANT: Rank, David R.  
; APPLICANT: Hanzel, David K.  
; APPLICANT: Chen, Wensheng  
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR  
; FILE REFERENCE: Aeomica-X-1  
; CURRENT APPLICATION NUMBER: US/09/864,761  
; CURRENT FILING DATE: 2001-05-23  
; PRIOR APPLICATION NUMBER: US 60/180,312  
; PRIOR FILING DATE: 2000-02-04  
; PRIOR APPLICATION NUMBER: US 60/207,456  
; PRIOR FILING DATE: 2000-05-26  
; PRIOR APPLICATION NUMBER: US 09/632,366  
; PRIOR FILING DATE: 2000-08-03  
; PRIOR APPLICATION NUMBER: GB 24263.6  
; PRIOR FILING DATE: 2000-10-04  
; PRIOR APPLICATION NUMBER: US 60/236,359  
; PRIOR FILING DATE: 2000-09-27  
; PRIOR APPLICATION NUMBER: PCT/US01/00666  
; PRIOR FILING DATE: 2001-01-30  
; PRIOR APPLICATION NUMBER: PCT/US01/00667  
; PRIOR FILING DATE: 2001-01-30  
; PRIOR APPLICATION NUMBER: PCT/US01/00664  
; PRIOR FILING DATE: 2001-01-30  
; PRIOR APPLICATION NUMBER: PCT/US01/00669  
; PRIOR FILING DATE: 2001-01-30  
; PRIOR APPLICATION NUMBER: PCT/US01/00665  
; PRIOR FILING DATE: 2001-01-30  
; PRIOR APPLICATION NUMBER: PCT/US01/00668  
; PRIOR FILING DATE: 2001-01-30  
; PRIOR APPLICATION NUMBER: PCT/US01/00663  
; PRIOR FILING DATE: 2001-01-30  
; PRIOR APPLICATION NUMBER: PCT/US01/00662  
; PRIOR FILING DATE: 2001-01-30  
; PRIOR APPLICATION NUMBER: PCT/US01/00661

; PRIOR FILING DATE: 2001-01-30  
; PRIOR APPLICATION NUMBER: PCT/US01/00670  
; PRIOR FILING DATE: 2001-01-30  
; PRIOR APPLICATION NUMBER: US 60/234,687  
; PRIOR FILING DATE: 2000-09-21  
; PRIOR APPLICATION NUMBER: US 09/608,408  
; PRIOR FILING DATE: 2000-06-30  
; PRIOR APPLICATION NUMBER: US 09/774,203  
; PRIOR FILING DATE: 2001-01-29  
; NUMBER OF SEQ ID NOS: 49117  
; SOFTWARE: Annomax Sequence Listing Engine vers. 1.1  
; SEQ ID NO 36764  
; LENGTH: 57  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
; FEATURE:  
; OTHER INFORMATION: MAP TO AL034406.1  
; OTHER INFORMATION: EXPRESSED IN BT474, SIGNAL = 0.6  
; OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 0.65  
; OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 0.74  
; OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 0.64  
; OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 0.7  
; OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 0.87  
; OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 0.73  
; OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 0.79  
; OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 0.84  
; OTHER INFORMATION: EST\_HUMAN HIT: AUI18230.1, EVALUE 3.00e-07  
US-09-864-761-36764

Query Match 57.1%; Score 24; DB 9; Length 57;  
Best Local Similarity 71.4%; Pred. No. 8e+02;  
Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 1 NVLMAMN 7  
:|:|:|:  
Db 21 SVAMAMN 27

## RESULT 43

US-10-044-359-2  
; Sequence 2, Application US/10044359  
; Publication No. US20020160454A1  
; GENERAL INFORMATION:  
; APPLICANT: Heriman, Rafael  
; APPLICANT: Wong, James F.  
; APPLICANT: Lee, Jian-Ming  
; TITLE OF INVENTION: SCORPION TOXINS  
; FILE REFERENCE: BB1367 US NA  
; CURRENT APPLICATION NUMBER: US/10/044,359  
; CURRENT FILING DATE: 2002-01-11  
; PRIOR APPLICATION NUMBER: 09/599,416  
; PRIOR FILING DATE: 2000-06-22  
; PRIOR APPLICATION NUMBER: 60/140,227  
; PRIOR FILING DATE: 1999-06-22  
; NUMBER OF SEQ ID NOS: 28  
; SOFTWARE: Microsoft Office 97  
; SEQ ID NO 2  
; LENGTH: 58  
; TYPE: PRT  
; ORGANISM: Hottentotta judaica  
US-10-044-359-2

Query Match 57.1%; Score 24; DB 14; Length 58;  
Best Local Similarity 50.0%; Pred. No. 8.1e+02;  
Matches 4; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

Qy 2 VLMAMNII 9  
:|:|:|:  
Db 10 IVFALNII 17

## RESULT 44

US-10-044-359-10

Tue Feb 17 17 11:55:47 2004

```
; Sequence 10, Application US/10044359
; Publication No. US20020160454A1
; GENERAL INFORMATION:
; APPLICANT: Heriman, Rafael
; APPLICANT: Wong, James F.
; APPLICANT: Lee, Jian-Ming
; TITLE OF INVENTION: SCORPION TOXINS
; FILE REFERENCE: HB1367 US NA
; CURRENT APPLICATION NUMBER: US/10/044,359
; CURRENT FILING DATE: 2002-01-11
; PRIOR APPLICATION NUMBER: 09/599,416
; PRIOR FILING DATE: 2000-06-22
; PRIOR APPLICATION NUMBER: 60/140,227
; PRIOR FILING DATE: 1999-06-22
; NUMBER OF SEQ ID NOS: 28
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 10
; LENGTH: 59
; TYPE: PRT
; ORGANISM: Hottentotta judaica
; US-10-044-359-10

Query Match 57.1%; Score 24; DB 14; Length 59;
Best Local Similarity 37.5%; Pred. No. 8.3e+02;
Matches 3; Conservative 4; Mismatches 1; Indels 0; Gaps 0;

Qy 2 VLMAMNII 9
Db 12 IVLVMNVI 19

RESULT 45
US-09-864-408A-72
; Sequence 72, Application US/09864408A
; Publication No. US20040009474A1
; GENERAL INFORMATION:
; APPLICANT: Leach, Martin D.
; APPLICANT: Shimkets, Richard A.
; TITLE OF INVENTION: No. US20040009474A1 Human Polynucleotides and Polypeptides Encod
; FILE REFERENCE: 21402-012
; CURRENT APPLICATION NUMBER: US/09/864,408A
; CURRENT FILING DATE: 2001-05-24
; PRIOR APPLICATION NUMBER: 60/206,690
; PRIOR FILING DATE: 2000-05-24
; NUMBER OF SEQ ID NOS: 9088
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 72
; LENGTH: 62
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (5)..(5)
; OTHER INFORMATION: Wherein Xaa may be any naturally occurring amino acid
US-09-864-408A-72

Query Match 57.1%; Score 24; DB 12; Length 62;
Best Local Similarity 62.5%; Pred. No. 8.7e+02;
Matches 5; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Qy 2 VLMAMNII 9
Db 12 VLRAINLI 19

Search completed: February 17, 2004, 11:02:20
Job time : 14.1881 secs
```



GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: February 17, 2004, 10:50:13 ; Search time 5.43564 Seconds  
(without alignments)  
159.230 Million cell updates/sec

Title: US-09-900-147-2

Perfect score: 42

Sequence: 1 NVLMAMNII 9

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283308 seqs, 96168682 residues

Total number of hits satisfying chosen parameters: 37673

Minimum DB seq length: 0

Maximum DB seq length: 100

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : PIR 76:\*

1: piri:\*

2: pir2:\*

3: pir3:\*

4: pir4:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	29	69.0	70	2 F89891	truncated transpos
2	29	69.0	98	2 F89915	truncated transpos
3	28	66.7	98	2 T17143	NADH2 dehydrogenas
4	28	66.7	98	2 T17174	NADH2 dehydrogenas
5	28	66.7	98	2 T17177	NADH2 dehydrogenas
6	28	66.7	98	2 T17162	NADH2 dehydrogenas
7	27	64.3	81	2 G91230	hypothetical prote
8	27	64.3	81	2 F86077	hypothetical prote
9	27	64.3	99	2 F90284	hypothetical prote
10	26	61.9	53	2 D91064	hypothetical prote
11	26	61.9	73	2 AH0720	probable membrane
12	26	61.9	91	2 T11163	NADH2 dehydrogenas
13	26	61.9	92	2 F90842	probable membrane
14	26	61.9	92	2 C85700	probable membrane
15	26	61.9	92	2 B64867	probable membrane
16	26	61.9	96	1 F71843	cytochrome c553 pr
17	26	61.9	98	2 AG3416	hypothetical prote
18	25	59.5	59	2 E84008	hypothetical prote
19	25	59.5	63	2 E64024	hypothetical prote
20	25	59.5	74	2 D89831	hypothetical prote
21	25	59.5	79	2 F95210	conserved domain p
22	25	59.5	87	2 S39696	ycwE protein - Bac
23	25	59.5	98	2 S41828	NADH2 dehydrogenas
24	25	59.5	98	2 T11449	NADH2 dehydrogenas
25	25	59.5	100	2 S53994	TodR protein - Pse
26	24	57.1	48	2 H82409	hypothetical prote
27	24	57.1	57	2 FN0624	alpha-internexin -
28	24	57.1	57	2 E82667	hypothetical prote
29	24	57.1	59	2 T16386	hypothetical prote

30 24 57.1 67 2 E82163 hypothetical prote  
31 24 57.1 68 2 AF2809 hypothetical prote  
32 24 57.1 70 2 I83374 protein-tyrosine k  
33 24 57.1 70 2 AI2538 hypothetical prote  
34 24 57.1 77 2 G82881 acyl carrier prote  
35 24 57.1 81 2 S40833 hypothetical 9.4K  
36 24 57.1 85 2 A69985 hypothetical prote  
37 24 57.1 86 2 B84222 hypothetical prote  
38 24 57.1 87 2 A55571 chloride conductan  
39 24 57.1 88 2 T17418 hypothetical prote  
40 24 57.1 89 2 G69034 hypothetical prote  
41 24 57.1 93 2 B39378 exo-alpha-sialidas  
42 24 57.1 93 2 G84065 hypothetical prote  
43 24 57.1 96 2 S00119 hypothetical prote  
44 24 57.1 98 2 S47878 NADH2 dehydrogenas  
45 24 57.1 98 2 S24920 NADH2 dehydrogenas

#### ALIGNMENTS

##### RESULT 1

F89891 truncated transposase [imported] - Staphylococcus aureus (strain N315)  
C;Species: Staphylococcus aureus  
C;Date: 10-May-2001 #sequence\_revision 10-May-2001 #text\_change 22-Oct-2001  
C;Accession: F89891  
R;Kuroda, M.; Ohta, T.; Uchiyama, I.; Baba, T.; Yuzawa, H.; Kobayashi, I.; Cui, L.; Oguch  
ma, A.; Mizutani-Ui, Y.; Hattori, M.; Ogasawara, N.; Hayashi, H.; Hiramatsu, K.  
C.; Shiba, T.; Hattori, M.; Ogasawara, N.; Hayashi, H.; Hiramatsu, K.  
Lancet 357, 1225-1240, 2001  
A;Title: Whole genome sequencing of methicillin-resistant Staphylococcus aureus.  
A;Reference number: A89758; MUID:21311952; PMID:11418146  
A;Accession: F89891  
A;Status: preliminary  
A;Molecule type: DNA  
A;Residues: 1-70 <KUR>  
A;Cross-references: GB:BA000018; PID:gl3700994; PIDN:BAB42290.1; GSPDB:GN00149  
A;Experimental source: strain N315  
C;Genetics:  
A;Gene: truncated(tnp)

Query Match 69.0%; Score 29; DB 2; Length 70;  
Best Local Similarity 85.7%; Pred. No. 22;  
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 2 VLMAMNII 8  
DB 22 VLMAMNII 28

##### RESULT 2

F89915 truncated transposase [imported] - Staphylococcus aureus (strain N315)  
C;Species: Staphylococcus aureus  
C;Date: 10-May-2001 #sequence\_revision 10-May-2001 #text\_change 22-Oct-2001  
C;Accession: F89915  
R;Kuroda, M.; Ohta, T.; Uchiyama, I.; Baba, T.; Yuzawa, H.; Kobayashi, I.; Cui, L.; Oguch  
ma, A.; Mizutani-Ui, Y.; Hattori, M.; Ogasawara, N.; Hayashi, H.; Hiramatsu, K.  
C.; Shiba, T.; Hattori, M.; Ogasawara, N.; Hayashi, H.; Hiramatsu, K.  
Lancet 357, 1225-1240, 2001  
A;Title: Whole genome sequencing of methicillin-resistant Staphylococcus aureus.  
A;Reference number: A89758; MUID:21311952; PMID:11418146  
A;Accession: F89915  
A;Status: preliminary  
A;Molecule type: DNA  
A;Residues: 1-98 <KUR>  
A;Cross-references: GB:BA000018; PID:gl3701187; PIDN:BAB42482.1; GSPDB:GN00149  
A;Experimental source: strain N315  
C;Genetics:  
A;Gene: truncated-SA

Query Match 69.0%; Score 29; DB 2; Length 98;

Tue Feb 17 11:55:47 2004

us-09-900-147-2.rpr

Query Match 66.7%; Score 28; DB 2; Length 98;  
Best Local Similarity 62.5%; Pred. No. 52;  
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 NVLMAMNI 8  
| | | | |  
DB 44 NTLMALNM 51

RESULT 3  
T17143  
NADH2 dehydrogenase (ubiquinone) (EC 1.6.5.3) chain 4L - Pygathrix bieti mitochondrion  
C:Species: Pygathrix bieti mitochondrion  
C:Date: 04-Mar-2000 #sequence\_revision 04-Mar-2000 #text\_change 03-Jun-2002  
C:Accession: T17143; T17165; T17168  
R:Wang, W.; Forstner, M.R.J.; Zhang, Y.P.; Lui, Z.M.; Wei, Y.; Huang, H.Q.; Hu, H.G.; Xi  
Int. J. Primatol. 18, 305-320, 1997  
A:Title: A phylogeny of Chinese leaf monkeys using mitochondrial ND3-ND4 gene sequences.  
A:Reference number: Z18709  
A:Accession: T17143  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-98 <W>  
A:Cross-references: EMBL:U92951; NID:g2290417; PID:g2290419; PIDN:AAD04678.1  
A:Experimental source: isolate DJ1  
A:Accession: T17165  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-98 <W>  
A:Cross-references: EMBL:U92957; NID:g2290441; PID:g2290443; PIDN:AAD08825.1  
A:Experimental source: isolate DJ4  
A:Accession: T17168  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-98 <W>  
A:Cross-references: EMBL:U92958; NID:g2290445; PID:g2290447; PIDN:AAD08828.1  
A:Experimental source: isolate DJ5  
A:Superfamily: NADH dehydrogenase (ubiquinone) chain 4L  
C:Keywords: membrane-associated complex; mitochondrion; NAD; oxidative phosphorylation;

Query Match 66.7%; Score 28; DB 2; Length 98;  
Best Local Similarity 62.5%; Pred. No. 52;  
Matches 5; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 NVLMAMNI 8  
| | | | |  
DB 44 NTLMALNM 51

RESULT 4  
T17174  
NADH2 dehydrogenase (ubiquinone) (EC 1.6.5.3) chain 4L - Pygathrix roxellana mitochondrion  
C:Species: Pygathrix roxellana mitochondrion  
C:Date: 04-Mar-2000 #sequence\_revision 04-Mar-2000 #text\_change 03-Jun-2002  
C:Accession: T17174; T17180  
R:Wang, W.; Forstner, M.R.J.; Zhang, Y.P.; Lui, Z.M.; Wei, Y.; Huang, H.Q.; Hu, H.G.; Xi  
Int. J. Primatol. 18, 305-320, 1997  
A:Title: A phylogeny of Chinese leaf monkeys using mitochondrial ND3-ND4 gene sequences.  
A:Reference number: Z18709  
A:Accession: T17174  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-98 <W>  
A:Cross-references: EMBL:U92961; NID:g2290457; PID:g2290459; PIDN:AAD04660.1  
A:Experimental source: isolate CUL  
A:Accession: T17180  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-98 <W>  
A:Cross-references: EMBL:U92963; NID:g2290465; PID:g2290467; PIDN:AAD04666.1  
A:Experimental source: isolate CUL3  
A:Superfamily: NADH dehydrogenase (ubiquinone) chain 4L  
C:Keywords: membrane-associated complex; mitochondrion; NAD; oxidative phosphorylation;

Query Match 66.7%; Score 28; DB 2; Length 98;  
Best Local Similarity 62.5%; Pred. No. 52;  
Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 NVLMAMNI 8  
| | | | |  
DB 44 NTLMALNM 51

RESULT 5  
T17177  
NADH2 dehydrogenase (ubiquinone) (EC 1.6.5.3) chain 4L [similarity] - Pygathrix roxellana  
C:Species: mitochondrion Pygathrix roxellana  
C:Date: 19-May-2000 #sequence\_revision 19-May-2000 #text\_change 03-Jun-2002  
C:Accession: T17177  
R:Wang, W.; Forstner, M.R.J.; Zhang, Y.P.; Lui, Z.M.; Wei, Y.; Huang, H.Q.; Hu, H.G.; Xi  
Int. J. Primatol. 18, 305-320, 1997  
A:Title: A phylogeny of Chinese leaf monkeys using mitochondrial ND3-ND4 gene sequences.  
A:Reference number: Z18709  
A:Accession: T17177  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-98 <W>  
A:Cross-references: EMBL:U92962; NID:g2290461; PID:g2290463; PIDN:AAD04663.1  
A:Superfamily: NADH dehydrogenase (ubiquinone) chain 4L  
C:Keywords: membrane-associated complex; mitochondrion; NAD; oxidative phosphorylation;

Query Match 66.7%; Score 28; DB 2; Length 98;  
Best Local Similarity 62.5%; Pred. No. 52;  
Matches 5; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 NVLMAMNI 8  
| | | | |  
DB 44 NTLMALNM 51

RESULT 6  
T17162  
NADH2 dehydrogenase (ubiquinone) (EC 1.6.5.3) chain 4L [similarity] - Pygathrix bieti  
C:Species: mitochondrion Pygathrix bieti  
C:Date: 19-May-2000 #sequence\_revision 19-May-2000 #text\_change 03-Jun-2002  
C:Accession: T17162  
R:Wang, W.; Forstner, M.R.J.; Zhang, Y.P.; Lui, Z.M.; Wei, Y.; Huang, H.Q.; Hu, H.G.; Xi  
Int. J. Primatol. 18, 305-320, 1997  
A:Title: A phylogeny of Chinese leaf monkeys using mitochondrial ND3-ND4 gene sequences.  
A:Reference number: Z18709  
A:Accession: T17162  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-98 <W>  
A:Cross-references: EMBL:U92956; NID:g2290437; PID:g2290439; PIDN:AAD08822.1  
A:Superfamily: NADH dehydrogenase (ubiquinone) chain 4L  
C:Keywords: membrane-associated complex; mitochondrion; NAD; oxidative phosphorylation;

Query Match 66.7%; Score 28; DB 2; Length 98;  
Best Local Similarity 62.5%; Pred. No. 52;  
Matches 5; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 NVLMAMNI 8  
| | | | |  
DB 44 NTLMALNM 51

RESULT 7  
G91230  
hypothetical protein Ecs4815 [imported] - Escherichia coli (strain O157:H7, substrain R11  
C:Species: Escherichia coli  
C:Date: 18-Jul-2001 #sequence\_revision 18-Jul-2001 #text\_change 18-Jul-2001  
C:Accession: G91230  
R:Hayashi, T.; Makino, K.; Ohnishi, M.; Kurokawa, K.; Ishii, K.; Yokoyama, K.; Han, C.G.  
gasawara, N.; Yasunaga, T.; Kuhara, S.; Shiba, T.; Hattori, M.; Shinagawa, H.  
DNA Res. 8, 11-22, 2001  
A:Title: Complete genome sequence of enterohemorrhagic Escherichia coli O157:H7 and genom

A:Reference number: A99629; MUID:21156231; PMID:11258796

A:Accession: G91230

A>Status: preliminary

A:Molecule type: DNA

A:Residues: 1-81 <HAY>

A:Cross-references: GB:BA000007; PIDN:BA838238.1; PID:gl3364291; GSPDB:GN00154

A:Experimental source: strain O157:H7, substrain RMD 050952

C:Genetics:

A:Gene: ECs4815

Query Match 64.3%; Score 27; DB 2; Length 81;

Best Local Similarity 62.5%; Pred. No. 70;

Matches 5; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 2 VLMAMNII 9

Db 8 VVMAMTV 15

RESULT 8

F86077 hypothetical protein yjiE [imported] - Escherichia coli (strain O157:H7, substrain EDL93)

C:Species: Escherichia coli

C>Date: 16-Feb-2001 #sequence\_revision 16-Feb-2001 #text\_change 14-Sep-2001

C:Accession: F86077

R:Perna, N.T.; Plunkett III, G.; Burland, V.; Mau, B.; Glasner, J.D.; Rose, D.J.; Mayhew

iller, L.; Grobeck, E.J.; Davis, N.W.; Lim, A.; Dimalanta, E.; Potamousis, K.; Apodaca,

Nature 409, 529-533, 2001

A>Title: Genome sequence of enterohemorrhagic Escherichia coli O157:H7.

A:Reference number: A85480; MUID:21074935; PMID:11206551

A:Accession: F86077

A>Status: preliminary

A:Molecule type: DNA

A:Residues: 1-81 <STO>

A:Cross-references: GB:AE005174; NID:gl3518781; PIDN:AAG59082.1; GSPDB:GN00145; UNGP:254

A:Experimental source: strain O157:H7, substrain EDL933

C:Genetics:

A:Gene: yjiE

Query Match 64.3%; Score 27; DB 2; Length 81;

Best Local Similarity 62.5%; Pred. No. 70;

Matches 5; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 2 VLMAMNII 9

Db 8 VVMAMTV 15

RESULT 9

F90284 hypothetical protein SSO1296 [imported] - Sulfolobus solfataricus

C:Species: Sulfolobus solfataricus

C>Date: 24-May-2001 #sequence\_revision 24-May-2001 #text\_change 24-May-2001

C:Accession: F90284

R;She, Q.; Singh, R.K.; Confalonieri, P.; Zivanovic, Y.; Allard, G.; Aways, M.J.; Chan-

gong, I.; Jeffries, A.C.; Kozera, C.J.; Medina, N.; Peng, X.; Thi-Ngoc, H.P.; Redder, H

arrett, R.A.; Ragan, M.A.; Sensen, C.W.; Van der Oost, J.

submitted to GenBank, April 2001

A:Description: Sulfolobus solfataricus complete genome.

A:Reference number: A99139

A:Accession: F90284

A>Status: preliminary

A:Molecule type: DNA

A:Residues: 1-99 <KUR>

A:Cross-references: GB:AE006641; NID:gl3814497; PIDN:AAK41533.1; GSPDB:GN00155

C:Genetics:

A:Gene: SSO1296

Query Match 64.3%; Score 27; DB 2; Length 99;

Best Local Similarity 44.4%; Pred. No. 88;

Matches 4; Conservative 4; Mismatches 1; Indels 0; Gaps 0;

QY 1 NVLMAMNII 9

Db 75 NIIISMNAI 83

RESULT 10

D91064

hypothetical protein ECs3484 [imported] - Escherichia coli (strain O157:H7, substrain RIN

C:Species: Escherichia coli

C>Date: 18-Jul-2001 #sequence\_revision 18-Jul-2001 #text\_change 18-Jul-2001

C:Accession: D91064

R;Hayashi, T.; Makino, K.; Ohnishi, M.; Kurokawa, K.; Ishii, K.; Yokoyama, K.; Han, C.G.;

gasawara, N.; Yasunaga, T.; Kuhara, S.; Shiba, T.; Hattori, M.; Shinagawa, H.

DNA Res. 8, 11-22, 2001

A>Title: Complete genome sequence of enterohemorrhagic Escherichia coli O157:H7 and genom

A:Reference number: A99629; MUID:21156231; PMID:11258796

A:Accession: D91064

A>Status: preliminary

A:Molecule type: DNA

A:Residues: 1-53 <HAY>

A:Cross-references: GB:BA000007; PIDN:BA836907.1; PID:gl3362955; GSPDB:GN00154

A:Experimental source: strain O157:H7, substrain RMD 050952

C:Genetics:

A:Gene: ECs3484

Query Match 61.9%; Score 26; DB 2; Length 53;

Best Local Similarity 50.0%; Pred. No. 73;

Matches 4; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 1 NVLMAMNII 8

Db 43 NIVAAAMNL 50

RESULT 11

AH0720

probable membrane protein STY1908 [imported] - Salmonella enterica subsp. enterica serova

C:Species: Salmonella enterica subsp. enterica serovar Typhi

A>Note: this species has also been called Salmonella typhi

C>Date: 09-Nov-2001 #sequence\_revision 09-Nov-2001 #text\_change 18-Nov-2002

C:Accession: AH0720

R;Parkhill, J.; Dougan, G.; James, K.D.; Thomson, N.R.; Pickard, D.; Wain, J.; Churcher,

th, T.; Connor, P.; Cronin, A.; Davis, P.; Davies, R.M.; Dowd, L.; White, N.; Farrar,

S.; Moule, S.; O'Gaora, P.

Nature 413, 848-852, 2001

A:Authors: Parry, C.; Quail, M.; Rutherford, K.; Simmonds, M.; Skelton, J.; Stevens, K.;

A>Title: Complete genome sequence of a multiple drug resistant Salmonella enterica serova

A:Reference number: AB0502; MUID:21534947; PMID:11677608

A:Accession: AH0720

A>Status: preliminary

A:Molecule type: DNA

A:Residues: 1-73 <PAR>

A:Cross-references: GB:AL513382; PIDN:CAD02137.1; PID:gl6502971; GSPDB:GN00176

C:Genetics:

A:Gene: STY1908

Query Match 61.9%; Score 26; DB 2; Length 73;

Best Local Similarity 44.4%; Pred. No. 1e+02;

Matches 4; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 1 NVLMAMNII 9

Db 11 NVLMGLGLV 19

RESULT 12

T11163

NADH2 dehydrogenase (ubiquinone) (EC 1.6.5.3) chain 4L - hardbacked tick (Rhipicephalus

C:Species: mitochondrion Rhipicephalus sanguineus

C>Date: 16-Jul-1999 #sequence\_revision 16-Jul-1999 #text\_change 03-Jun-2002

C:Accession: T11163

R;Black IV, W.C.; Roehrdanz, R.L.

Mol. Biol. Evol. 15, 1772-1785, 1998

A>Title: Mitochondrial gene order is not conserved in arthropods: prostrate and metastr

A;Reference number: Z17252; MUID:99083443; PMID:9866211

A;Accession: T1163

A;Status: preliminary; translated from GB/EMBL/DBJ

A;Molecule type: DNA

A;Residues: 1-91 <BLA>

A;Cross-references: EMBL:AF081829; NID:94164556; PID:94164567; PIDN:AAD05527.1

C;Genetics:

A;Genome: mitochondrion

C;Superfamily: NADH dehydrogenase (ubiquinone) chain 4L

C;Keywords: membrane-associated complex; mitochondrion; NAD; oxidative phosphorylation;

Query Match 61.9%; Score 26; DB 2; Length 91;

Best Local Similarity 50.0%; Pred. No. 1.3e+02;

Matches 4; Conservative 4; Mismatches 0; Indels 0; Gaps 0;

QY 2 VLMAMNII 9

DB 82 LNMAMNLI 89

RESULT 13

F90842 Probable membrane protein ychH - Escherichia coli (strain O157:H7, substrain RIMD 050995)

C;Species: Escherichia coli

C;Date: 18-Jul-2001 #sequence\_revision 18-Jul-2001 #text\_change 27-Nov-2001

C;Accession: F90842

R;Hayaishi, T.; Makino, K.; Ohnishi, M.; Kurokawa, K.; Ishii, K.; Yokoyama, K.; Han, C.G.

GASAWARA, N.; Yasunaga, T.; Kuhara, S.; Shiba, T.; Hattori, M.; Shinagawa, H.

Nature Res. 8, 11-22, 2001

A;Title: Complete genome sequence of enterohemorrhagic Escherichia coli O157:H7 and gen

A;Reference number: A99629; MUID:21156231; PMID:11258796

A;Accession: F90842

A;Status: preliminary

A;Molecule type: DNA

A;Residues: 1-92 <HAY>

A;Cross-references: GB:BA000007; PIDN:BA035133.1; PID:g13361175; GSPDB:GN00154

A;Experimental source: strain O157:H7, substrain RIMD 050952

C;Genetics:

A;Gene: ECs1710

Query Match 61.9%; Score 26; DB 2; Length 92;

Best Local Similarity 44.4%; Pred. No. 1.3e+02;

Matches 4; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 1 NVLMAMNII 9

DB 11 NVLMGLGLV 19

RESULT 14

C85700 Probable membrane protein ychH - Escherichia coli (strain O157:H7, substrain EDL933)

C;Species: Escherichia coli

C;Date: 16-Feb-2001 #sequence\_revision 16-Feb-2001 #text\_change 27-Nov-2001

C;Accession: C85700

R;Perna, N.T.; Plunkett III, G.; Burland, V.; Mau, B.; Glasner, J.D.; Rose, D.J.; Mayhew

iller, L.; Grotbeck, E.J.; Davis, N.W.; Lim, A.; Dimalanta, E.; Potamouis, K.; Apodaca,

Nature 409, 529-533, 2001

A;Title: Genome sequence of enterohemorrhagic Escherichia coli O157:H7.

A;Reference number: A85480; MUID:21074935; PMID:11206551

A;Accession: C85700

A;Status: preliminary

A;Molecule type: DNA

A;Residues: 1-92 <STO>

A;Cross-references: GB:AE005174; NID:g12514918; PIDN:AAG56063.1; GSPDB:GN00145; UWGP:Z19

A;Experimental source: strain O157:H7, substrain EDL933

C;Genetics:

A;Gene: ychH

Query Match 61.9%; Score 26; DB 2; Length 92;

Best Local Similarity 44.4%; Pred. No. 1.3e+02;

Matches 4; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 1 NVLMAMNII 9

DB 11 NVLMGLGLV 19

RESULT 15

B64867

Probable membrane protein ychH - Escherichia coli (strain K-12)

C;Species: Escherichia coli

C;Date: 12-Sep-1997 #sequence\_revision 17-Sep-1997 #text\_change 01-Mar-2002

C;Accession: B64867

R;Blattner, F.R.; Plunkett III, G.; Bloch, C.A.; Perna, N.T.; Burland, V.; Riley, M.; Col

A.; Rose, D.J.; Mau, B.; Shao, Y.

Science 277, 1453-1462, 1997

A;Title: The complete genome sequence of Escherichia coli K-12.

A;Reference number: A64720; MUID:97426617; PMID:9278503

A;Accession: B64867

A;Status: nucleic acid sequence not shown; translation not shown

A;Molecule type: DNA

A;Residues: 1-92 <BLAT>

A;Cross-references: GB:AE000219; GB:U00096; NID:g1787453; PIDN:AAC74289.1; PID:g1787456;

A;Experimental source: strain K-12, substrain MG1655

C;Genetics:

A;Gene: ychH

C;Keywords: transmembrane protein

F;9-25/Domain: transmembrane #status predicted <TM1>

F;45-61/Domain: transmembrane #status predicted <TM2>

Query Match 61.9%; Score 26; DB 2; Length 92;

Best Local Similarity 44.4%; Pred. No. 1.3e+02;

Matches 4; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 1 NVLMAMNII 9

DB 11 NVLMGLGLV 19

RESULT 16

F71843

Cytochrome c553 precursor - Helicobacter pylori (strain J99)

C;Species: Helicobacter pylori

A;Variety: strain J99

C;Date: 03-Dec-1999 #sequence\_revision 03-Dec-1999 #text\_change 03-Mar-2000

C;Accession: F71843

R;Alm, R.A.; Ling, L.S.L.; Moir, D.T.; King, B.L.; Brown, E.D.; Doig, P.C.; Smith, D.R.;

Ives, C.; Gibson, R.; Merberg, D.; Mills, S.D.; Jiang, Q.; Taylor, D.E.; Vovis, G.F.;

Nature 397, 176-180, 1999

A;Title: Genomic sequence comparison of two unrelated isolates of the human gastric patho

A;Reference number: A71800; MUID:99120557; PMID:9923682

A;Accession: F71843

A;Molecule type: DNA

A;Residues: 1-96 <ARN>

A;Cross-references: GB:AE001542; GB:AE001439; NID:g4155739; PIDN:AAD06721.1; PID:g415574;

A;Experimental source: strain J99

C;Genetics:

A;Gene: jhp1148

C;Superfamily: cytochrome c6; cytochrome c6 homology

C;Keywords: chromoprotein; electron transfer; heme; iron; metalloprotein

F;1-19/Domain: signal sequence #status predicted <SIG>

F;20-96/Product: cytochrome c553 #status predicted <MAT>

F;20-92/Domain: cytochrome c6 homology <CYC>

F;29,32/Binding site: heme (Cys) (covalent) #status predicted

F;33/Binding site: heme iron (His) (axial ligand) #status predicted

Query Match 61.9%; Score 26; DB 1; Length 96;

Best Local Similarity 62.5%; Pred. No. 1.4e+02;

Matches 5; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 NVLMAMNI 8

DB 15 NVLMATDV 22



```

A:Start codon: GTG
C:Superfamily: NADH dehydrogenase (ubiquinone) chain 4L
C:Keywords: membrane-associated complex; mitochondrion; NAD; oxidative phosphorylation;
Query Match 59.5%; Score 25; DB 2; Length 98;
Best Local Similarity 62.5%; Pred. No. 2.4e+02;
Matches 5; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 NVLMAMNI 8
DB 7 NVLMAFSM 14
|||||::
|::|::|

RESULT 24
Tl1449
NADH2 dehydrogenase (ubiquinone) (EC 1.6.5.3) chain 4L - nine-banded armadillo mitochondrion
C:Species: mitochondrion Dasyapus novemcinctus (nine-banded armadillo)
C:Date: 16-Jul-1999 #sequence_revision 16-Jul-1999 #text_change 03-Jun-2002
C:Accession: Tl1449
R:Arnason, U.; Gullberg, A.; Janke, A.
Mol. Biol. Evol. 14, 762-768, 1997
A:Title: Phylogenetic analysis of mitochondrial DNA suggest a sister group relationship
A:Reference number: Z17272; MUID:97357423; PMID:9214749
A:Accession: Tl1449
A:Status: preliminary; translated from GB/EMBL/DBDJ
A:Molecule type: DNA
A:Residues: 1-98 <ARN>
A:Cross-references: EMBL:Y11832; NID:g2252500; PIDN:CAA72524.1; PID:g2252509
C:Genetics:
A:Genome: mitochondrion
A:Genetic code: SGC1
A:Note: NADH4L
C:Superfamily: NADH dehydrogenase (ubiquinone) chain 4L
C:Keywords: membrane-associated complex; mitochondrion; NAD; oxidative phosphorylation;
Query Match 59.5%; Score 25; DB 2; Length 98;
Best Local Similarity 50.0%; Pred. No. 2.4e+02;
Matches 4; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 1 NVLMAMNI 8
DB 7 NIIIMAFSI 14
|::|::|
|::|::|

RESULT 25
TODR protein - Pseudomonas putida
C:Species: Pseudomonas putida
C:Date: 28-Oct-1996 #sequence_revision 13-Mar-1997 #text_change 08-Oct-1999
C:Accession: S53994
R:Wang, Y.; Rawlings, M.; Gibson, D.T.; Labbe, D.; Bergeron, H.; Brousseau, R.; Lau, P.C.
Mol. Gen. Genet. 246, 570-579, 1995
A:Title: Identification of a membrane protein and a truncated LysR-type regulator associ
A:Reference number: S53993; MUID:95214617; PMID:7535376
A:Accession: S53994
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-100 <WAN>
A:Cross-references: GB:U18304; NID:ig747633; PIDN:AA43317.1; PID:g747635
A:Note: the authors translated the codon GAG for residue 71 as Asp and CGC for residue 8
Query Match 59.5%; Score 25; DB 2; Length 100;
Best Local Similarity 44.4%; Pred. No. 2.4e+02;
Matches 1; Indels 0; Gaps 0;

```

```

C;Species: Vibrio cholerae
C;Date: 18-Aug-2000 #sequence_revision 20-Aug-2000 #text_change 02-Feb-2001
C;Accession: H82409
R;Heidelberg, J.F.; Eisen, J.A.; Nelson, W.C.; Clayton, R.A.; Gwinn, M.L.; Dodson, R.J.;
chardson, D.; Ermolaeva, M.D.; Vamathevan, J.; Bacs, S.; Qin, H.; Dragoi, I.; Sellers, P.
1, R.R.; Mekalanos, J.J.; Venter, J.C.; Fraser, C.M.
Nature 406, 477-483, 2000
A;Title: DNA Sequence of both chromosomes of the cholera pathogen Vibrio cholerae.
A;Reference number: A82035; MUID:20406833; PMID:10952301
A;Accession: H82409
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-48 <HEI>
A;Cross-references: GB:AE004412; GB:AE003853; NID:g9658269; PIDN:AAF96742.1; GSPDB:GN001
A;Experimental source: serogroup O1; strain N16961; biotype El Tor
C;Genetics:
A;Gene: VCA0844
A;Map position: 2

Query Match 57.1%; Score 24; DB 2; Length 48;
Best Local Similarity 50.0%; Pred. No. 1.8e+02;
Matches 4; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 2 VLMAMNII 9
DB 36 VMVKMNIV 43

RESULT 27
PN0624
alpha-internexin - pig (fragments)
C;Species: Sus scrofa domestica (domestic pig)
C;Date: 05-Aug-1994 #sequence_revision 05-Aug-1994 #text_change 17-Mar-1999
C;Accession: PN0624
R;Tanaka, J.; Ogawara, M.; Ando, S.; Shibata, M.; Yatani, R.; Kusagawa, M.; Inagaki, M.
Biochem. Biophys. Res. Commun. 196, 115-123, 1993
A;Title: Phosphorylation of a 62kD porcine alpha-internexin, a newly identified intermed
A;Reference number: PN0624; MUID:94029981; PMID:8216281
A;Accession: PN0624
A;Molecule type: protein
A;Residues: 1-18;19-28;29-45;46-57 <TAN>
A;Comment: This intermediate filament protein polymerizes to form filament itself, in vi
C;Superfamily: cytoskeletal keratin
C;Keywords: brain; intermediate filament

Query Match 57.1%; Score 24; DB 2; Length 57;
Best Local Similarity 62.5%; Pred. No. 2.2e+02;
Matches 5; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 NVLMAMNI 8
DB 43 NVKMWLDI 50

RESULT 28
E82667
hypothetical protein XF1541 [imported] - Xylella fastidiosa (strain 9a5c)
C;Species: Xylella fastidiosa
C;Date: 18-Aug-2000 #sequence_revision 20-Aug-2000 #text_change 20-Aug-2000
C;Accession: E82667
R;Anonymous, The Xylella fastidiosa Consortium of the Organization for Nucleotide Sequen
Nature 406, 151-157, 2000
A;Title: The genome sequence of the plant pathogen Xylella fastidiosa.
A;Reference number: A82515; MUID:20365717; PMID:10910347
A;Note: for a complete list of authors see reference number A59328 below
A;Accession: E82667
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-57 <SIM>
A;Cross-references: GB:AE003984; GB:AE003849; NID:g9106581; PIDN:AAF84350.1; GSPDB:GN001
A;Experimental source: strain 9a5c
R;Simpson, A.J.G.; Reinach, F.C.; Arruda, P.; Abreu, F.A.; Acencio, M.; Alvarenga, R.; A
Briones, M.R.S.; Bueno, M.R.P.; Camargo, A.A.; Camargo, L.E.A.; Carraro, D.M.; Carrer, H

```

```

as-Neto, E.; Docena, C.; El-Dorry, H.; Pacincani, A.P.; Ferreira, A.J.S.
submitted to GenBank, June 2000
A;Authors: Ferreira, V.C.A.; Ferro, E.L.; Kitajima, J.P.; Krieger, J.E.; Kuramae, E.E.; Laigre
J.D.; Junqueira, M.L.; Kemper, E.L.; Madeira, H.M.F.; Marino, C.L.; Marques, M.V.; Martins, E.
chado, M.A.; Madeira, A.M.B.N.; Madeira, A.Y.; Menck, C.F.M.; Miracca, E.C.; Miyaki, C.Y.;
A;Authors: Martins, E.M.F.; Matsukuma, A.Y.; Menck, C.F.M.; de Oliveira, R.C.; Palmieri, D.A.
R.G.; Nunes, L.R.; Oliveira, M.A.; de Rosa Jr., V.E.; de Sa, R.G.; Santelli, R.V.; Sawasaki
Rodrigues, V.; Rosa, A.J. de M.; de Rosa Jr., V.E.; da Silva, A.M.; Silva Jr., W.A.; da Silveira
M.; Tshuko, M.H.; Vallada, H.; Van Sluys, M.A.; Verjovski-Almeida, S.; Vettore, A.L.; Z
A;Reference number: A59328
A;Contents: annotation
C;Genetics:
A;Gene: XF1541

Query Match 57.1%; Score 24; DB 2; Length 57;
Best Local Similarity 33.3%; Pred. No. 2.2e+02;
Matches 3; Conservative 5; Mismatches 1; Indels 0; Gaps 0;

QY 1 NVLMAMNII 9
DB 37 NLIAAINVL 45

RESULT 29
T16386
hypothetical protein F47B7.1 - Caenorhabditis elegans
C;Species: Caenorhabditis elegans
C;Date: 20-Sep-1999 #sequence_revision 20-Sep-1999 #text_change 20-Sep-1999
C;Accession: T16386
R;Leimbach, D.
submitted to the EMBL Data Library, October 1995
A;Description: The sequence of C. elegans cosmid F47B7.
A;Reference number: Z18504
A;Accession: T16386
A;Status: preliminary; translated from GB/EMBL/DD8J
A;Molecule type: DNA
A;Residues: 1-59 <LEI>
A;Cross-references: EMBL:U39646; NID:g1049351; PID:g1049358; PIDN:AAA80373.1; CESP:F47B7
C;Genetics:
A;Gene: CESP:F47B7.1
A;Introns: 18/3

Query Match 57.1%; Score 24; DB 2; Length 59;
Best Local Similarity 55.6%; Pred. No. 2.3e+02;
Matches 5; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 1 NVLMAMNII 9
DB 30 NMHVAVNII 38

RESULT 30
E82163
hypothetical protein VC1733 [imported] - Vibrio cholerae (strain N16961 serogroup O1)
C;Species: Vibrio cholerae
C;Date: 18-Aug-2000 #sequence_revision 20-Aug-2000 #text_change 02-Feb-2001
C;Accession: E82163
R;Heidelberg, J.F.; Eisen, J.A.; Nelson, W.C.; Clayton, R.A.; Gwinn, M.L.; Dodson, R.J.;
chardson, D.; Ermolaeva, M.D.; Vamathevan, J.; Bacs, S.; Qin, H.; Dragoi, I.; Sellers, P.
1, R.R.; Mekalanos, J.J.; Venter, J.C.; Fraser, C.M.
Nature 406, 477-483, 2000
A;Title: DNA Sequence of both chromosomes of the cholera pathogen Vibrio cholerae.
A;Reference number: A82035; MUID:20406833; PMID:10952301
A;Accession: E82163
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-67 <HEI>
A;Cross-references: GB:AE004251; GB:AE003852; NID:g9656248; PIDN:AAF94883.1; GSPDB:GN001
A;Experimental source: serogroup O1; strain N16961; biotype El Tor
C;Genetics:
A;Gene: VC1733
A;Map position: 1

```

Tue Feb 17 11:55:47 2004

```

AI2538
hypothetical protein asl7596 [imported] - Nostoc sp. (strain PCC 7120) plasmid pCC7120bet.
C;Species: Nostoc sp. PCC 7120
A;Note: Nostoc sp. strain PCC 7120 is a synonym of Anabaena sp. strain PCC 7120
C;Date: 14-Dec-2001 #sequence_revision 14-Dec-2001 #text_change 09-Dec-2002
C;Accession: AI2538
R;Kaneko, T.; Nakamura, Y.; Wolk, C.P.; Kuritz, T.; Sasamoto, S.; Watanabe, A.; Iriguchi,
Nakazaki, N.; Shimpo, S.; Sugimoto, M.; Takazawa, M.; Yamada, M.; Yasuda, M.; Tabata, S
DNA Res. 8, 205-213, 2001
A;Title: Complete Genomic Sequence of the Filamentous Nitrogen-fixing Cyanobacterium Anati
A;Reference number: AB1807; MUID:21595285; PMID:11759840
A;Accession: AI2538
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-70 <KUR>
A;Cross-references: GB:AP003602; PIDN:BAB77239.1; PID:gl7134681; GSPDB:GN00181
A;Experimental source: strain PCC 7120
C;Genetics:
A;Gene: asl7596
A;Genome: plasmid

Query Match 57.1%; Score 24; DB 2; Length 70;
Best Local Similarity 44.4%; Pred. No. 2.7e+02;
Matches 4; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 1 NVLMAMNII 9
| | | | |
Db 28 NLLQIINIV 36

RESULT 34
GB2881
acyl carrier protein UU506 [imported] - Ureaplasma urealyticum
C;Species: Ureaplasma urealyticum
C;Date: 18-Aug-2000 #sequence_revision 20-Aug-2000 #text_change 20-Aug-2000
C;Accession: GB2881
R;Glass, J.I.; Lefkowitz, E.J.; Glass, J.S.; Heiner, C.R.; Chen, E.Y.; Casseil, G.H.
submitted to GenBank, February 2000
A;Description: The complete sequence of Ureaplasma urealyticum: Alternate views of a min
A;Reference number: AB2870
A;Accession: GB2881
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-77 <GLA>
A;Cross-references: GB:AE002149; GB:AF222894; NID:96899503; PIDN:AAF30918.1; GSPDB:GN001
A;Experimental source: serovar 3; biovar 1
C;Genetics:
A;Gene: acpp; UU506
A;Genetic Code: SGC3

Query Match 57.1%; Score 24; DB 2; Length 77;
Best Local Similarity 71.4%; Pred. No. 3e+02;
Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 3 LMAMNII 9
| | | | |
Db 37 LSAANLI 43

RESULT 35
S40833
hypothetical 9.4K protein (glnA-fdhE intergenic region) - Escherichia coli (strain K-12)
N;Alternate names: hypothetical protein o81
C;Species: Escherichia coli
C;Date: 19-May-1994 #sequence_revision 01-Sep-1995 #text_change 01-Mar-2002
C;Accession: S40833; D65194
R;Plunkett III, G.; Burland, V.; Daniels, D.L.; Blattner, F.R.
Nucleic Acids Res. 21, 3391-3398, 1993
A;Title: Analysis of the Escherichia coli genome. III. DNA sequence of the region from 8,
A;Reference number: S40802; MUID:93347969; PMID:8346018
A;Accession: S40833
A;Status: preliminary; nucleic acid sequence not shown; translation not shown
A;Molecule type: DNA

```



A;Residues: 1-81 <PLU>  
A;Cross-references: EMBL:LI9201; NID:G304961; PIDN:AAB03022.1; PID:G304993  
A;Note: The nucleotide sequence was submitted to the EMBL Data Library, October 1993  
R;Blattner, F.R.; Plunkett III, G.; Bloch, C.A.; Perna, N.T.; Burland, V.; Riley, M.; C  
A.; Rose, D.J.; Mau, B.; Shao, Y.  
Science 277, 1453-1462, 1997  
A;Title: The complete genome sequence of *Escherichia coli* K-12.  
A;Reference number: A64720; MUID:97426617; PMID:9278503  
A;Accession: D65194  
A;Status: preliminary; nucleic acid sequence not shown; translation not shown  
A;Molecule type: DNA  
A;Residues: 1-81 <BLAT>  
A;Cross-references: GB:AE000464; GB:U00096; NID:G92367324; PIDN:AAD13451.1; PID:gl790322;  
A;Experimental source: strain K-12, substrain MG1655  
C;Genetics:  
A;Gene: yjiE

Query Match 57.1%; Score 24; DB 2; Length 81;  
Best Local Similarity 50.0%; Pred. No. 3.2e+02;  
Matches 4; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 2 VLMAMNII 9  
:|||||:  
Db 8 LVWAMTV 15

RESULT 36  
A69985  
hypothetical protein yeha - *Bacillus subtilis*  
C;Species: *Bacillus subtilis*  
C;Date: 05-Dec-1997 #sequence\_revision 05-Dec-1997 #text\_change 15-Oct-1999  
C;Accession: A69985  
R;Kunst, F.; Ogasawara, N.; Moszer, I.; Albertini, A.M.; Alloni, G.; Azevedo, V.; Ber  
C.; Bron, S.; Brouillet, S.; Bruschi, C.V.; Caldwell, B.; Capuano, V.; Carter, N.M.; Ch  
A.; Ehrlich, S.D.; Emerson, P.T.; Entian, K.D.; Errington, J.; Fabret, C.; Ferrari, E.  
Nature 390, 249-256, 1997  
A;Authors: Foulger, D.; Fritz, C.; Fujita, M.; Fujita, Y.; Fuma, S.; Galizzi, A.; Gall  
iech, J.; Harwood, C.R.; Henaut, A.; Hilbert, H.; Holsappel, S.; Hosono, S.; Hullo, M.F.  
Koester, P.; Konigstein, G.; Krogh, S.; Kumano, M.; Kurita, K.; Lapidus, A.; Lardinois,  
A;Authors: Lauber, J.; Lazarevic, V.; Lee, S.M.; Levine, A.; Liu, H.; Masuda, S.; Maues  
Y. M.; Ogawa, K.; Ogiwara, A.; Oudega, B.; Park, S.H.; Parro, V.; Pohl, T.M.; Portetel  
Rieger, M.; Rivolt, C.; Rocha, E.; Roche, B.; Rose, M.; Sadate, Y.; Sato, T.; Scanlon,  
A;Authors: Schleich, S.; Schroeter, R.; Scoffone, P.; Sekiguchi, J.; Sekowska, A.; Seron  
akeuchi, M.; Tamakoshi, A.; Tanaka, T.; Terpstra, P.; Tognoni, A.; Tosato, V.; Uchiyama,  
T.; Winters, P.; Wipat, A.; Yamamoto, K.; Yamane, K.; Yasumoto, K.; Yata, K.; Yoshida, K  
A;Authors: Yoshikawa, H.F.; Zumbstein, E.; Yoshikawa, H.; Danchin, A.  
A;Title: The complete genome sequence of the Gram-positive bacterium *Bacillus subtilis*.  
A;Reference number: A69580; MUID:98044033; PMID:9384377  
A;Accession: A69985  
A;Status: preliminary; nucleic acid sequence not shown; translation not shown  
A;Molecule type: DNA  
A;Residues: 1-85 <KUN>  
A;Cross-references: GB:Z99118; GB:AL009126; NID:G2635200; PIDN:CAB14821.1; PID:el184110;  
A;Experimental source: strain 168  
C;Genetics:  
A;Gene: yehA

Query Match 57.1%; Score 24; DB 2; Length 85;  
Best Local Similarity 50.0%; Pred. No. 3.4e+02;  
Matches 4; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 2 VLMAMNII 9  
:|||||:  
Db 57 VLTAVNV 64

RESULT 37  
B84222  
hypothetical protein Vng0641c [imported] - *Halobacterium* sp. NRC-1  
C;Species: *Halobacterium* sp. NRC-1  
C;Date: 02-Feb-2001 #sequence\_revision 02-Feb-2001 #text\_change 02-Feb-2001  
C;Accession: B84222  
R;Ng, W.V.; Kennedy, S.P.; Mahairas, G.G.; Berquist, B.; Pan, M.; Shukla, H.D.; Lasky, S

; Leithauser, B.; Keller, K.; Cruz, R.; Danson, M.J.; Hough, D.W.; Maddocks, D.G.; Jablor  
Jung, K.H.; Alam, M.; Freitas, T.  
Proc. Natl. Acad. Sci. U.S.A. 97, 12176-12181, 2000  
A;Authors: Hou, S.; Daniels, C.J.; Dennis, P.P.; Omer, A.D.; Ebhardt, H.; Lowe, T.M.; Li  
A;Title: Genome sequence of *Halobacterium* species NRC-1.  
A;Reference number: A84160; MUID:20504483; PMID:11016950  
A;Accession: B84222  
A;Status: preliminary  
A;Molecule type: DNA  
A;Residues: 1-86 <STO>  
A;Cross-references: GB:AE004437; NID:G10580232; PIDN:AAG19142.1; GSPDB:GN00138  
C;Genetics:  
A;Gene: VNG0641C

Query Match 57.1%; Score 24; DB 2; Length 86;  
Best Local Similarity 71.4%; Pred. No. 3.4e+02;  
Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 3 LAMAMNII 9  
:|||||:  
Db 54 LMAAQIL 60

RESULT 38  
A55571  
chloride conductance inducer Mat-8 - human  
C;Species: *Homo sapiens* (man)  
C;Date: 23-Mar-1995 #sequence\_revision 23-Mar-1995 #text\_change 28-May-1999  
C;Accession: A55571  
R;Morrison, B.W.; Moorman, J.R.; Kowdley, G.C.; Kobayashi, Y.M.; Jones, L.R.; Leder, P.  
J. Biol. Chem. 270, 2176-2182, 1995  
A;Title: Mat-8, a novel phospholemman-like protein expressed in human breast tumors, ind  
A;Reference number: A55571; MUID:95138184; PMID:7836447  
A;Accession: A55571  
A;Status: preliminary  
A;Molecule type: mRNA  
A;Residues: 1-87 <MOR>  
A;Cross-references: GB:X93036; GB:S74645; NID:G1085025; PIDN:CAA63604.1; PID:e211793; PI  
C;Keywords: transmembrane protein

Query Match 57.1%; Score 24; DB 2; Length 87;  
Best Local Similarity 75.0%; Pred. No. 3.5e+02;  
Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 2 VLMAMNII 9  
:|||||:  
Db 47 VLCAMGII 54

RESULT 39  
T17418  
hypothetical protein pttL - *Pseudomonas fluorescens*  
C;Species: *Pseudomonas fluorescens*  
C;Date: 15-Oct-1999 #sequence\_revision 15-Oct-1999 #text\_change 15-Oct-1999  
C;Accession: T17418  
R;Nowak-Thompson, B.; Chaney, N.; Wing, J.S.; Gould, S.J.; Lopez, J.E.  
J. Bacteriol. 181, 2166-2174, 1999  
A;Title: Characterization of the pyoluteorin biosynthetic gene cluster of *Pseudomonas fl*  
A;Reference number: Z18776; MUID:99194726; PMID:10094695  
A;Accession: T17418  
A;Status: preliminary; translated from GB/EMBL/DDJ  
A;Molecule type: DNA  
A;Residues: 1-88 <NOW>  
A;Cross-references: EMBL:AF081920; NID:G4582974; PID:G4582982; PIDN:AAD24885.1  
C;Genetics:  
A;Gene: pttL

Query Match 57.1%; Score 24; DB 2; Length 88;  
Best Local Similarity 50.0%; Pred. No. 3.5e+02;  
Matches 4; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 2 VLMAMNII 9  
:|||||:

Db 39 ILNSMNV 46

RESULT 40

G69034

hypothetical protein MTH1255 - Methanobacterium thermoautotrophicum (strain Delta H)

C:Species: Methanobacterium thermoautotrophicum

C>Date: 05-Dec-1997 #sequence\_revision 05-Dec-1997 #text\_change 22-Oct-1999

C:Accession: G69034

R:Smith, D.R.; Doucette-Stamm, L.A.; Deloughery, C.; Lee, H.; Dubois, J.; Aldredge, T.; Qiu, D.; Spadafora, R.; Vicaire, R.; Wang, Y.; Wierzbowski, J.; Gibson, R.; Jiواني, N.; ki, S.; Church, G.M.; Daniels, C.J.; Mao, J.; Rice, F.; Noelling, J.; Reeve, J.N.

J. Bacteriol. 179, 7135-7155, 1997

A:Title: Complete genome sequence of Methanobacterium thermoautotrophicum Delta H: func

A:Reference number: A69000; MUID:98037514; PMID:9371463

A:Accession: G69034

A>Status: preliminary; nucleic acid sequence not shown; translation not shown

A:Molecule type: DNA

A:Residues: 1-89 <MTH>

A:Cross-references: GB:AE000892; GB:AE000666; NID:g2622364; PIDN:AAB85744.1; PID:g262236

A:Experimental source: strain Delta H

C:Genetics:

A:Gene: MTH1255

Query Match 57.1%; Score 24; DB 2; Length 89;

Best Local Similarity 50.0%; Pred. No. 3.5e+02;

Matches 4; Conservative 4; Mismatches 0; Indels 0; Gaps 0;

Qy 2 VLMMNII 9

Db 47 LLLAVNII 54

RESULT 41

B39378

exo-alpha-sialidase (EC 3.2.1.18) SA85-1.2 - Trypanosoma cruzi (fragment)

C:Species: Trypanosoma cruzi

C>Date: 20-Mar-1992 #sequence\_revision 20-Mar-1992 #text\_change 11-Jan-2000

C:Accession: B39378

R:Kahn, S.; Colbert, T.G.; Wallace, J.C.; Hoegland, N.A.; Eisen, H.

Proc. Natl. Acad. Sci. U.S.A. 88, 4481-4485, 1991

A:Title: The major 85-kDa surface antigen of the mammalian-stage forms of Trypanosoma ch

A:Reference number: A39378; MUID:91239592; PMID:2034687

A:Accession: B39378

A>Status: preliminary

A:Molecule type: DNA

A:Residues: 1-93 <KAH>

A:Cross-references: GB:M62736

C:Superfamily: trypanostigote-specific surface antigen

C:Keywords: glycosidase; hydrolase

Query Match 57.1%; Score 24; DB 2; Length 93;

Best Local Similarity 100.0%; Pred. No. 3.7e+02;

Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 3 LMMN 7

Db 15 LMMN 19

RESULT 42

G84065

hypothetical protein BH3327 [imported] - Bacillus halodurans (strain C-125)

C:Species: Bacillus halodurans

C>Date: 01-Dec-2000 #sequence\_revision 01-Dec-2000 #text\_change 15-Jun-2001

C:Accession: G84065

R:Takami, H.; Nakasone, K.; Takaki, Y.; Maeno, G.; Sasaki, R.; Masui, N.; Fuji, F.; Hira

Nucleic Acids Res. 28, 4317-4331, 2000

A:Title: Complete genome sequence of the alkaliphilic bacterium Bacillus halodurans and

A:Reference number: A83650; MUID:20512582; PMID:11058132

A:Accession: G84065

A>Status: preliminary

A:Molecule type: DNA

Query Match 57.1%; Score 24; DB 2; Length 93;

Best Local Similarity 100.0%; Pred. No. 3.7e+02;

Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 3 LMMN 7

Db 15 LMMN 19

RESULT 43

S00119

hypothetical protein C - Proteus vulgaris

C:Species: Proteus vulgaris

C>Date: 31-Dec-1990 #sequence\_revision 31-Dec-1990 #text\_change 14-Apr-2003

C:Accession: S00119

R:Coile, S.T.

Eur. J. Biochem. 167, 481-488, 1987

A:Title: Nucleotide sequence and comparative analysis of the frd operon encoding the funa

d-linked ampC cephalosporinase gene.

A:Reference number: S00107; MUID:88004470; PMID:3308458

A:Accession: S00119

A:Molecule type: DNA

A:Residues: 1-96 <COL>

A:Cross-references: EMBL:X06151; NID:g45907; PIDN:CAA29511.1; PID:g45910

C:Superfamily: [NiFe]-hydrogenase maturation chaperone

Query Match 57.1%; Score 24; DB 2; Length 96;

Best Local Similarity 83.3%; Pred. No. 3.8e+02;

Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 NVLMAM 6

Db 70 NALMAM 75

RESULT 44

S47878

NADH2 dehydrogenase (ubiquinone) (EC 1.6.5.3) chain 4L - North American opossum mitochondr

C:Species: mitochondrion Didelphis virginiana, Didelphis marsupialis virginiana (North A

C>Date: 19-Mar-1997 #sequence\_revision 19-Mar-1997 #text\_change 03-Jun-2002

C:Accession: S47878; S42818

R:Janke, A.; Feldmaier-Fuchs, G.; Thomas, W.K.; von Haeseler, A.; Paasebo, S.

Genetics 137, 243-256, 1994

A:Title: The marsupial mitochondrial genome and the evolution of placental mammals.

A:Reference number: S47870; MUID:94333786; PMID:8056314

A:Accession: S47878

A:Molecule type: DNA

A:Residues: 1-98 <JAN>

A:Cross-references: EMBL:Z29573; NID:g452251; PIDN:CAA82685.1; PID:g555588

A:Experimental source: liver

C:Genetics:

A:Gene: NADH4L

A:Genome: mitochondrion

A:Genetic code: SGC1

C:Superfamily: NADH dehydrogenase (ubiquinone) chain 4L

C:Keywords: membrane-associated complex; mitochondrion; NAD; oxidative phosphorylation;

Query Match 57.1%; Score 24; DB 2; Length 98;

Best Local Similarity 62.5%; Pred. No. 3.9e+02;

Matches 5; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Qy 2 VLMMNII 9

Db 2 VLKLNII 9

RESULT 45

S24920  
NADH2 dehydrogenase (ubiquinone) (EC 1.6.5.3) chain 4L - finback whale mitochondrion  
C:Species: mitochondrion Balaenoptera physalus (finback whale, common rorqual)  
C>Date: 31-Dec-1993 #sequence\_revision 31-Dec-1993 #text\_change 03-Jun-2002  
C:Accession: A58851; S24920  
R:Arnason, U.; Gullberg, A.; Widegren, B.  
J. Mol. Evol. 33, 556-568, 1991  
A:Title: The complete nucleotide sequence of the mitochondrial DNA of the fin whale, Bal  
A:Reference number: A58850; MUID:92139449; PMID:1779436  
A:Accession: A58851  
A:Molecule type: DNA  
A:Residues: 1-98 <ARN>  
A:Cross-references: GB:X61145; NID:gl2772; PIDN:CRA43447.1; PID:gl2779  
A:Note: submitted to the EMBL Data Library, August 1991  
C:Genetics:  
A:Genome: mitochondrion  
A:Genetic code: SGCI  
C:Superfamily: NADH dehydrogenase (ubiquinone) chain 4L  
C:Keywords: membrane-associated complex; mitochondrion; NAD; oxidative phosphorylation;  
Query Match 57.1%; Score 24; DB 2; Length 98;  
Best Local Similarity 50.0%; Pred. No. 3.9e+02;  
Matches 4; Conservative 3; Mismatches 1; Indels 0; Gaps 0;  
QY 1 NVLMAMNI 8  
Db |:|:|:| :|:  
7 NILMAFSM 14

Search completed: February 17, 2004, 10:58:14  
Job time : 6.43564 secs

**THIS PAGE BLANK (USPTO)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: February 17, 2004, 10:50:13 ; Search time 2.85149 Seconds  
(without alignments)  
148.428 Million cell updates/sec

Title: US-09-900-147-2  
Perfect score: 42  
Sequence: 1 NULMANNIII 9

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 127863 seqs, 47026705 residues

Total number of hits satisfying chosen parameters: 13973

Minimum DB seq length: 0  
Maximum DB seq length: 100

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : SwissProt\_41:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	26	61.9	69	1 Y010_TROWT	P59486 tropheryma
2	26	61.9	91	1 NULM_RHISA	O99826 thipicephal
3	26	61.9	92	1 YCHH_ECOLI	P31807 escherichia
4	26	61.9	96	1 C553_HELPJ	O32129 helicobacte
5	25	59.5	56	1 SCP2_MESMA	Q9njp7 mesobuthus
6	25	59.5	56	1 SCP3_MESMA	Q9ubd1 mesobuthus
7	25	59.5	87	1 YWCE_BACSU	P39603 bacillus su
8	25	59.5	98	1 NULM_BALMU	P41301 balaenopter
9	25	59.5	98	1 NULM_DASNO	O21333 dasypus nov
10	24	57.1	59	1 YV31_CAREL	O20516 caenorhabdi
11	24	57.1	77	1 ACPH_UREPA	Q9ppv4 ureaplasma
12	24	57.1	87	1 FAX3_HUMAN	Q14802 homo sapien
13	24	57.1	96	1 YFRC_PROVU	P20927 proteus vul
14	24	57.1	98	1 NULM_BALPH	P24976 balaenopter
15	24	57.1	98	1 NULM_DIDMA	P41307 didelphis m
16	23	54.8	48	1 ATP8_CANPA	P17345 candida par
17	23	54.8	83	1 Y666_CHLTR	O84673 chlamydia t
18	23	54.8	84	1 Y710_CHLPN	Q9z7j5 chlamydia p
19	23	54.8	87	1 RPOH_THEAC	Q03588 thermoplas
20	23	54.8	91	1 RL31_HALN1	Q9hmo0 halobacteri
21	23	54.8	91	1 VAPD_HAEIN	P71351 haemophilus
22	23	54.8	95	1 Y736_ARCFU	O29522 archoeoglob
23	23	54.8	98	1 NULM_BOVIN	P03902 bos taurus
24	23	54.8	98	1 NULM_SHEEP	O8754 ovis aries
25	23	54.8	98	1 RL3E_METTH	O27127 methanobact
26	23	54.8	99	1 YF16_METJA	Q58911 methanococc
27	22	52.4	32	1 ADHR_DROYA	P28487 drosophila
28	22	52.4	78	1 YJZJ_ECOLI	P55914 escherichia
29	22	52.4	82	1 C2_OXNO	P05526 oxytricha n
30	22	52.4	82	1 S61B_ARATH	P38389 arabidopsis
31	22	52.4	86	1 PAKI_SYNY3	P27212 synechocyst
32	22	52.4	91	1 Y069_AERPE	Q9y9f0 aeropyrum p
33	22	52.4	92	1 YC23_METJA	O58620 methanococc

34	22	52.4	96	1 C553_HELPY	O25825 helicobacte
35	22	52.4	96	1 Y143_BORBU	O51168 borrelia bu
36	22	52.4	98	1 NULM_HIPAM	Q922v3 hippopotamu
37	22	52.4	98	1 NULM_HYLLA	Q95709 hylobates l
38	22	52.4	98	1 NULM_PIG	P56632 sus scrofa
39	22	52.4	98	1 NULM_PISOC	P25000 pisaster oc
40	22	52.4	99	1 Y233_MYCGE	P47475 mycoplasma
41	21	50.0	53	1 YBAM_ECOLI	P45807 escherichia
42	21	50.0	56	1 Y546_METJA	Q57966 methanococc
43	21	50.0	65	1 YC33_CVAPA	P48273 cyanophora
44	21	50.0	70	1 CSPA_RICCN	Q92gv1 rickettsia
45	21	50.0	70	1 CSPA_RICPR	Q9zcp9 rickettsia

## ALIGNMENTS

## RESULT 1

Y010\_TROWT STANDARD; PRT; 69 AA.  
AC P59486;  
DT 15-SRP-2003 (Rel. 42, Created)  
DT 15-SRP-2003 (Rel. 42, Last sequence update)  
DT 15-SRP-2003 (Rel. 42, Last annotation update)  
DE Hypothetical UPF0233 membrane protein TWT010/TW010.  
GN TWT010 OR TW010.  
OS Tropheryma whippelii (strain Twist) (Whipple's bacillus), and  
OS Tropheryma whippelii (strain TW08/27) (Whipple's bacillus).  
OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;  
OC Micrococineae; Cellulomonadaceae; Tropheryma.  
OX NCBI\_TaxID=203267, 218496,  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=Twist / Genotype 2A;  
RA Racult D., Audic S., Robert C., Ogata H., Suhre K., Drancourt M.,  
RA Claverie J.-M.;  
RT "Tropheryma whippelii illustrates the diversity of gene loss patterns  
in small genome bacterial pathogens.";  
RL Submitted (JUL-2002) to the EMBL/GenBank/DBJ databases.  
[2]  
RN SEQUENCE FROM N.A.  
RC STRAIN=TW08/27;  
RA MEDLINE=22495039; PubMed=12606174;  
RA Bentley S.D., Maiwald M., Murphy L.D., Pallen M.J., Yeats C.A.,  
RA Dover L.G., Norbertczak H.T., Besra G.S., Quail M.A., Harris D.E.,  
RA von Herbay A., Goble A., Rutter S., Squares R., Squares S.,  
RA Barrell B.G., Parkhill J., Rellman D.A.;  
RT "Sequencing and analysis of the genome of the Whipple's disease  
bacterium Tropheryma whippelii.";  
RL Lancet 361:627-634(2003).  
CC -1- SUBCELLULAR LOCATION: Integral membrane protein (Potential).  
CC -1- SIMILARITY: Belongs to the UPF0233 family.  
CC  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
between the Swiss Institute of Bioinformatics and the EMBL outstation -  
the European Bioinformatics Institute. There are no restrictions on its  
use by non-profit institutions as long as its content is in no way  
modified and this statement is not removed. Usage by and for commercial  
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
or send an email to [license@sib-sib.ch](mailto:license@sib-sib.ch)).  
CC  
CC  
CC EMBL; AE016850; AAC044107.1; -;  
CC EMBL; BX251410; CAD66702.1; -;  
CC HAWAP; MF\_00631; -; 1.  
KW Hypothetical protein; Transmembrane.  
FT TRANSMEM 14 34 POTENTIAL.  
FT TRANSMEM 45 65 POTENTIAL.  
SQ SEQUENCE 69 AA; 7723 MW; 3C83E5B3F154DD24 CRC64;

Query Match 61.9%; Score 26; DB 1; Length 69;  
Best Local Similarity 55.6%; Pred. No. 46;  
Matches 5; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

Tue Feb 17 11:55:47 2004

us-09-900-147-2-rsp

QY 1 NVLMANNII 9  
 Db 50 NIIAFGII 58

RESULT 2  
 NULM RHISA STANDARD; PRT; 91 AA.  
 AC O99826;  
 DT 30-MAY-2000 (Rel. 39, Last sequence update)  
 DT 30-MAY-2000 (Rel. 39, Last sequence update)  
 DT 16-OCT-2001 (Rel. 40, Last annotation update)  
 DE ND4L.  
 GN ND4L.  
 OS Rhinoceros squalidus (Brown dog tick).  
 OG Mitochondrion.  
 OC Eukaryota; Metazoa; Arthropoda; Chelicerata; Arachnida; Acari;  
 OC Parasitiformes; Ixodida; Ixodidae; Rhinoceros.  
 OC NCBI\_TaxID=34632;  
 [1]  
 RN SEQUENCE FROM N.A.  
 RP MEDLINE=99083443; PubMed=9866211;  
 RX Black W.C. IV, Roehrdanz R.L.;  
 RA "Mitochondrial gene order is not conserved in arthropods: prostriate  
 RT and metastriate tick mitochondrial genomes.";  
 RL Mol. Biol. Evol. 15:1772-1785(1998).  
 CC -1- CATALYTIC ACTIVITY: NADH + ubiquinone = NAD(+) + ubiquinol.  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC

DR EMBL; AF081829; AAC05527.1; -;  
 DR PIR; T11163;  
 DR InterPro; IPR001133; Oxidored4L.  
 DR InterPro; IPR003214; Oxidred4L.  
 DR Pfam; PF00420; oxidored\_g2; 1.  
 DR ProDom; PD000359; Oxidred4L; 1.  
 KW Oxidoreductase; NAD; Ubiquinone; Mitochondrion.  
 SQ SEQUENCE 91 AA; 10598 MW; D2B51EC5F1D056EE CRC64;

Query Match 61.9%; Score 26; DB 1; Length 91;  
 Best Local Similarity 50.0%; Pred. NO. 62;  
 Matches 4; Conservative 4; Mismatches 0; Indels 0; Gaps 0;

QY 2 VLMANNII 9  
 Db 82 LMMSNLI 89

RESULT 3  
 YCHH\_ECOLI STANDARD; PRT; 92 AA.  
 AC F31807;  
 DT 01-JUL-1993 (Rel. 26, Created)  
 DT 01-JUL-1993 (Rel. 26, Last sequence update)  
 DT 28-FEB-2003 (Rel. 41, Last annotation update)  
 DE Hypothetical protein ychh.  
 GN YCHH OR B1205 OR C1663 OR Z1976 OR ECS1710.  
 OS Escherichia coli,  
 OS Escherichia coli O6, and  
 OS Escherichia coli O157:H7.  
 OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;  
 OC Enterobacteriaceae; Escherichia.  
 OC NCBI\_TaxID=562, 217992, 83334;  
 [1]  
 RN SEQUENCE FROM N.A.  
 RP STRAIN=K12;  
 RX MEDLINE=92007806; PubMed=1833189;

Garcia-Villegas M.R., de la Vega F.M., Galindo J.M., Segura M.,  
 Buckingham R.H., Guarneros G.;  
 "Peptidyl-tRNA hydrolase is involved in lambda inhibition of host  
 protein synthesis";  
 EMBO J. 10:3549-3555(1991).  
 [2]  
 RN REVISIONS.  
 RP STRAIN=K12;  
 RC MEDLINE=95129849; PubMed=7828865;  
 RX Galindo J.M., de la Vega F.M., Guarneros G.;  
 "Open reading frame flanking the peptidyl-tRNA hydrolase-encoding  
 RT gene of Escherichia coli";  
 RL Gene 151:153-156(1994).  
 [3]  
 RN SEQUENCE FROM N.A.  
 RP STRAIN=K12 / MG1655;  
 RC MEDLINE=97426617; PubMed=9278503;  
 RX Blattner F.R., Plunkett G. III, Bloch C.A., Perna N.T., Burland V.,  
 RA Riley M., Collado-Vides J., Glasner J.D., Rode C.K., Mayhew G.F.,  
 RA Gregor J., Davis N.W., Kirkpatrick H.A., Goeden M.A., Rose D.J.,  
 RA Mau B., Shao Y.;  
 RT "The complete genome sequence of Escherichia coli K-12";  
 RL Science 277:1453-1474(1997).  
 [4]  
 RN SEQUENCE FROM N.A.  
 RP STRAIN=K12;  
 RC MEDLINE=97061202; PubMed=8905232;  
 RX Oshima T., Aiba H., Baba T., Fujita K., Hayashi K., Honjo A.,  
 RA Ikemoto K., Inada T., Itoh T., Kajihara M., Kanai K., Kashimoto K.,  
 RA Kimura S., Kitagawa M., Makino K., Masuda S., Miki T., Mizobuchi K.,  
 RA Mori H., Motomura K., Nakamura Y., Nishimoto H., Nishio Y., Saito N.,  
 RA Sampei G., Seki Y., Tagami H., Takemoto K., Wada C., Yamamoto Y.,  
 RA Yano M., Horuchi T.;  
 RT "A 718-kb DNA sequence of the Escherichia coli K-12 genome  
 RT corresponding to the 12.7-28.0 min region on the linkage map";  
 RL DNA Res. 3:1137-155(1996).  
 [5]  
 RN SEQUENCE FROM N.A.  
 RP STRAIN=O6:H1 / CFT073 / ATCC 700928;  
 RC MEDLINE=22388234; PubMed=12471157;  
 RX Welch R.A., Burland V., Plunkett G. III, Redford P., Roesch P.,  
 RA Rasko D., Buckles E.L., Liou S.-R., Boutin A., Hackett J., Stroud D.,  
 RA Mayhew G.F., Rose D.J., Zhou S., Schwartz D.C., Perna N.T.,  
 RA Mobley H.L.T., Donnenberg M.S., Blattner F.R.;  
 RT "Extensive mosaic structure revealed by the complete genome sequence  
 RT of uropathogenic Escherichia coli";  
 RL Proc. Natl. Acad. Sci. U.S.A. 99:17020-17024(2002).  
 [6]  
 RN SEQUENCE FROM N.A.  
 RP STRAIN=O157:H7 / EDL933 / ATCC 700927;  
 RC MEDLINE=21074935; PubMed=11206551;  
 RX Perna N.T., Plunkett G. III, Burland V., Mau B., Glasner J.D.,  
 RA Rose D.J., Mayhew G.F., Evans P.S., Gregor J., Kirkpatrick H.A.,  
 RA Posfai G., Hackett J., Klink S., Boutin A., Shao Y., Miller L.,  
 RA Grobeck E.J., Davis N.W., Lim J., Yen G., Schwartz D.C.,  
 RA Apodaca J., Anantharaman T.S., Lin J., Perna N.T.,  
 RA Welch R.A., Blattner F.R.;  
 RT "Genome sequence of enterohaemorrhagic Escherichia coli O157:H7";  
 RL Nature 409:529-533(2001).  
 [7]  
 RN SEQUENCE FROM N.A.  
 RP STRAIN=O157:H7 / RIMD 0509952;  
 RC MEDLINE=21156231; PubMed=11258796;  
 RX Hayashi T., Makino K., Ohnishi M., Kurokawa K., Ishii K., Yokoyama K.,  
 RA Han C.-G., Ohtsubo E., Nakayama K., Murata T., Tanaka M., Tobe T.,  
 RA Iida T., Takami H., Honda T., Sasaki C., Ogasawara N., Yasunaga T.,  
 RA Kuhara S., Shiba T., Hattori M., Shinagawa H.;  
 RT "Complete genome sequence of enterohaemorrhagic Escherichia coli  
 RT O157:H7 and genomic comparison with a laboratory strain K-12";  
 RL DNA Res. 8:11-22(2001).  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -

CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

EMBL; X61941; CAA43947.1; -.  
EMBL; AF000219; AAC74289.1; -.  
EMBL; D90756; BAA35063.1; -.  
EMBL; AE016759; AAAN8028.1; -.  
EMBL; AE005338; AAC5068.1; -.  
EMBL; AP025566; BAB35133.1; -.  
PIR; B48667; B64867.  
PIR; C85700; C85700.  
PIR; F90842; F90842.  
ECogene; EG11533; wcbH.

RESULT 4	CS53_HELPJ	STANDARD;	PRT;	96 AA.
ID	-C553_HELPJ			
AC	Q9ZJZ9;			
DT	28-FEB-2003 (Rel. 41, Created)			
DT	28-FEB-2003 (Rel. 41, Last sequence update)			
DT	28-FEB-2003 (Rel. 41, Last annotation update)			
DE	Cytochrome C-553 precursor (C553).			
GN	JHP1148.			
OS	Helicobacter pylori J99 (Campylobacter pylori J99).			
OC	Bacteria; Proteobacteria; Epsilonproteobacteria; Campylobacteriales;			
OC	Helicobacteraceae; Helicobacter.			
OX	NCBI_TaxID=85963;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RK	MEDLINE=99120557; PubMed=9923692;			
RA	Alm R.A., Ling L.-S.L., Moir D.T., King B.L., Brown E.D., Doig P.C.,			
RA	Smith D.R., Noonan B., Guild B.C., deJonge B.L., Carmel G.,			
RA	Tummino P.J., Caruso A., Uria-Nickelsen M., Mills D.M., Ives C.,			
RA	Gibson R., Merberg D., Mills S.D., Jiang Q., Taylor D.E., Vovis G.F.,			
RA	Trust T.J.;			
RT	"Genomic sequence comparison of two unrelated isolates of the human			
RT	gastric pathogen Helicobacter pylori.";			
RL	Nature 397:176-180(1999).			
CC	-I- FUNCTION: Natural electron acceptor for a formate dehydrogenase.			
CC	-I- SUBCELLULAR LOCATION: Periplasmic.			
CC	-I- PTM: BINDS ONE HEME GROUP PER MOLECULE.			
CC	-I- SIMILARITY: BELONGS TO THE CLASS I CYTOCHROME C FAMILY.			

EMBL; AE001542; AAD06721.1; -.  
 PIR; F71843; F71843.  
 DR HSSP; P04032; 2DVH.  
 DR InterPro; IPR003088; Cyt CI.  
 DR InterPro; IPR002329; Cyt-CIC.  
 DR InterPro; IPR00345; Cyt-heme\_bind.  
 DR Pfam; PF00034; cytochrome c 1.  
 DR PRINTS; PR00605; CYTCHROME6IC.  
 DR

```

ProDom: PD004020; Cyt_C_bact; 1.
DR PROSITE; PS00190; CYTOCHROME C; 1.
DR Electron transport; Heme; Signal; Complete proteome.
FT SIGNAL 1 19 BY SIMILARITY.
FT CHAIN 20 96 CYTOCHROME C-553.
FT BINDING 29 29 HEME (COVALENT).
FT BINDING 32 32 HEME (COVALENT).
FT METAL 33 33 IRON (HEME AXIAL LIGAND) (BY
SIMILARITY).
FT METAL 73 73 IRON (HEME AXIAL LIGAND) (BY
SIMILARITY)
SQ SEQUENCE 96 AA; 10354 MW; 3E607AED5422AD82 CRC64;

Query Match 61.9%; Score 26; DB 1; Length 96;
Best Local Similarity 62.5%; Pred. No. 65;
Matches 5; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Qy 1 NVLMAMNI 8
Db 15 NVLMATDV 22

```

RESULT 5

SCP2_MESMA	STANDARD;	PRT;	56 AA.
ID	SCP2_MESMA		
AC	Q9NJF7; P58491;		
DT	28-FEB-2003 (Rel. 41, Created)		
DT	28-FEB-2003 (Rel. 41, Last sequence update)		
DT	15-SEP-2003 (Rel. 42, Last annotation update)		
DE	Neurotoxin Bmp02 precursor (Potassium ion channel blocker P02P).		
OS	Mesobuthus martensii (Mancurian scorpion)		
OC	Eukaryota; Metazoa; Arthropoda; Chelicerata; Arachnida; Scorpiones;		
OC	Buthoidea; Buthidae; Mesobuthus.		
XC	NCBI_TaxID=34549;		
RN	[1]		
RP	SEQUENCE FROM N.A.		
RC	TISSUE=Venom gland;		
RA	MEDLINE=99402983; PubMed=10471839;		
RX	Zhu S.-Y., Li W.-X., Zeng X.-C., Jiang D.-H., Mao X., Liu H.;		
RT	'Molecular cloning and sequencing of two 'short chain' and two 'long		
RT	chain' K(+) channel-blocking peptides from the Chinese scorpion Buthus		
RL	martensii Karsch.;"		
RL	FEBS Lett. 457:509-514 (1999).		
RN	[2]		
RP	STRUCTURE BY NMR OF 29-56.		
RC	TISSUE=Venom;		
RA	MEDLINE=205330297; PubMed=11076505;		
RX	Xu Y.-Q., Wu J.H., Pei J.-M., Shi Y.-Y., Ji Y.H., Tong Q.-C.;		
RT	'Solution structure of Bmp02, a new potassium channel blocker from the		
RT	venom of the Chinese scorpion Buthus martensii Karsch.;"		
RL	Biochemistry 39:13669-13675 (2000).		
CC	-1- FUNCTION: Blocks potassium channels.		
CC	-1- SUBCELLULAR LOCATION: Secreted.		
CC	-1- TISSUE SPECIFICITY: Expressed by the venom gland.		
CC	-1- SIMILARITY: BELONGS TO THE SCORPION LEIUTOXIN FAMILY.		

DR	EWBL; AF132975; RAFP1296.1; -.	
DR	PDB; 1DU9; 28-MAR-01.	
KW	Toxin; Neurotoxin; ionic channel inhibitor;	
KW	Potassium channel inhibitor; Signal; 3D-structure.	
FT	SIGNAL	1
FT	CHAIN	28
FT	CHAIN	29
FT	DISULFID	56
FT	DISULFID	31
FT	DISULFID	47
FT	DISULFID	34
FT	DISULFID	52
FT	DISULFID	38
FT	DISULFID	54
FT	NEUROTOXIN BMP02.	

DB	10	IVLAMNVM	17
RESULT 7			
YWCE_BACSU	STANDARD;	PRT;	87 AA.
ID_YWCE_BACSU			
AC P39603;	1995 (Rel. 31, Created)		
DT 01-FEB-1995	(Rel. 31, Last sequence update)		
DT 28-FEB-2003	(Rel. 41, Last annotation update)		
DE Hypothetical protein ywce precursor.			
GN YWCE OR IPA-41R.			
OS Bacillus subtilis.			
OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.			
OX NCBI_TaxID=1423;			
RN [1]			
RP SEQUENCE FROM N.A.			
RC STRAIN=168;			
RX MEDLINE=95020537; PubMed=7934828;			
RA Glaser P., Kunst F., Arnaud M., Coudart M.P., Gonzales W.,			
RA Hulio M.F., Ionescu M., Lubochinsky B., Marcelino L., Moser I.,			
RA Presecan E., Sencanu M., Schneider E., Schweizer J., Vertes A.,			
RA Rapoport G., Danchin A.;			
RT "Bacillus subtilis genome project: cloning and sequencing of the 97			
RL kb region from 325 degrees to 333 degrees.";			
RL Mol. Microbiol. 10:371-384(1993).			
RN [2]			
RP SEQUENCE FROM N.A.			
RC STRAIN=168;			
RX MEDLINE=98044033; PubMed=9384377;			
RA Kunst F., Ogasawara N., Moszer I., Albertini A.M., Alloni G.,			
RA Azevedo V., Bertero M.G., Bessieres P., Bolotin A., Borchert S.,			
RA Borriess R., Boursier L., Brans A., Braun M., Brignell S.C., Bron S.,			
RA Brouillet S., Bruschi C.V., Caldwell B., Capuano V., Carter N.M.,			
RA Choi S.K., Codani J.J., Connerton I.F., Cummings N.J., Daniel R.A.,			
RA Denizot F., Devine K.M., Dusterhoft A., Ehrlich S.D., Emmerson P.T.,			
RA Entian K.D., Errington J., Fabret C., Ferrazi A., Foulger D.,			
RA Fritz C., Fujita M., Fujita Y., Fuma S., Gallizzi A., Galleron N.,			
RA Ghim S.Y., Glaser P., Goffeau A., Golightly E.J., Grandi G.,			
RA Guelpi G., Guy B.J., Haga K., Haiech J., Harwood C.R., Henaut A.,			
RA Hilbert H., Holtsappel S., Hosono S., Hulio M.F., Itaya M., Jones L.,			
RA Joris B., Karamata D., Kaehara Y., Klaert-Blanchard M., Klein C.,			
RA Kobayashi Y., Koetter P., Koningsstein G., Krogh S., Kumano M.,			
RA Kurita K., Lapidus A., Lardinois S., Lauber J., Lazarevic V.,			
RA Lee S.M., Levine A., Liu H., Masuda S., Maue C., Medigue C.,			
RA Medina N., Mellado R.P., Mizuno M., Moestl D., Nakai S., Noback M.,			
RA Noone D., O'Reilly M., Ogawa K., Ogiwara A., Oudega B., Park S.H.,			
RA Parro V., Pohl T.M., Portetelle D., Porwollik S., Prescott A.M.,			
RA Pieper E., Pujic P., Purnelle B., Roche B., Rose M., Sadale Y.,			
RA Rieger M., Rivolta C., Rocha E., Roche B., Schreter R., Scifo B., Soldo B.,			
RA Sato T., Scanlan E., Schleich S., Seror S.J., Serrif P., Shin B.S.,			
RA Sekiguchi J., Sekowska A., Seror S.J., Takahashi H., Takemaru K.,			
RA Sorokin A., Tacconi E., Tagaki T., Terpatra P., Tognoni A.,			
RA Takeuchi M., Tamakoshi A., Tanaka T., Vannier F., Vassarotti A.,			
RA Tosato V., Uchiyama S., Vandenberg M., Vannier F., Weitzenecker T.,			
RA Viari A., Wambutt R., Wedler E., Wedler H., Weitzenecker T.,			
RA Winters P., Wipatt A., Yamamoto H., Yamane K., Yasumoto K., Yata K.,			
RA Yoshida K., Yoshikawa H.F., Zumbstein E., Yoshikawa H., Danchin A.;			
RT "The complete genome sequence of the Gram-positive bacterium Bacillus			
RT subtilis.";			
Nature 390:249-256(1997).			
RL -I- SUBCELLULAR LOCATION: Integral membrane protein (Potential).			
CC This SWISS-PROT entry is copyright. It is produced through a collaboration			
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation			
CC at the European Bioinformatics Institute. There are no restrictions on its			
CC use by non-profit institutions as long as its content is in no way			
CC modified and this statement is not removed. Usage by and for commercial			
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/			
CC or send an email to license@isb-sib.ch).			
CC -----			
DR EMBL; X73124; CAAS1597.1; -			
DR EMBL; Z99123; CAB15839.1; -			



```

DR PIR; S39696; S39696.
DR Subtilisin; BGI0587; YWCE.
KW Hypothetical protein; Transmembrane; Signal; Complete proteome.
FT SIGNAL 1 29 POTENTIAL.
FT CHAIN 30 87 HYPOTHETICAL PROTEIN YWCE.
FT TRANSMEM 30 46 POTENTIAL.
FT TRANSMEM 55 78 POTENTIAL.
SQ SEQUENCE 87 AA; 9959 MW; CC74135F423940CB CRC64;

Query Match 59.5%; Score 25; DB 1; Length 87;
Best Local Similarity 50.0%; Pred. No. 98;
Matches 4; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 2 VLMAMNII 9
Db 59 ILFAVNVII 66

RESULT 8
NULM_BALMU STANDARD; PRT; 98 AA.
AC P41301;
DT 01-FEB-1995 (Rel. 31, Created)
DT 01-FEB-1995 (Rel. 31, Last sequence update)
DT 01-NOV-1997 (Rel. 35, Last annotation update)
DE NADH-ubiquinone oxidoreductase chain 4L (EC 1.6.5.3).
GN MTND4L OR ND4L OR NADH4L.
OS Balaenoptera musculus (Blue whale).
OG Mitochondrion.
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Cetacea; Mysticeti;
OC Balaenopteridae; Balaenoptera.
OX NCBI_TaxID=9771;
RN 1;
RP SEQUENCE FROM N.A.
RX MEDLINE=94141932; PubMed=8308901;
RA Arnason U., Gullberg A.;
RT "Comparison between the complete mtDNA sequences of the blue and the
RL fin whale, two species that can hybridize in nature.";
J. Mol. Evol. 37:312-322(1993).
CC -|- CATALYTIC ACTIVITY: NADH + ubiquinone = NAD(+) + ubiquinol.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; X72204; CAA51003.1; -.
DR PIR; S41828; S41828.
DR InterPro; IPR001133; Oxidored 4L.
DR InterPro; IPR003214; Oxidred4L.
DR Pfam; PF00420; Oxidored_q2; 1.
DR ProDom; PD000359; Oxidred4L; 1.
KW Oxidoreductase; NAD; Ubiquinone; Mitochondrion.
OX NCBI_TaxID=9771;
RN 1;
RP SEQUENCE 98 AA; 10747 MW; 9F770651FE65ED1B CRC64;

Query Match 59.5%; Score 25; DB 1; Length 98;
Best Local Similarity 62.5%; Pred. No. 1.1e+02;
Matches 5; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 NVLMAMNI 8
Db 7 NVLMAFSM 14

RESULT 9
NULM_DASNO STANDARD; PRT; 98 AA.
AC O2133;
DT 15-JUL-1998 (Rel. 36, Created)
DT 15-JUL-1998 (Rel. 36, Last sequence update)
DT 15-JUL-1998 (Rel. 36, Last annotation update)
DE Hypothetical 6.8 kDa protein F47B7.1 in chromosome X.
GN F47B7.1.
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
OC Rhabditidae; Peloderinae; Caenorhabditis.
OX NCBI_TaxID=6239;
RN 1;
RP SEQUENCE FROM N.A.
RX STRAIN=Bristol N2;
RA Leimbach D.;
RT Submitted (NOV-1995) to the EMBL/GenBank/DBJ databases.
CC -|- SUBCELLULAR LOCATION: Integral membrane protein (Potential).
CC -|- SIMILARITY: BELONGS TO THE UPF0057 (PMP3) FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----

```

DR EMBL: U39646; AAA80373.1; -  
 DR PIR: T16386; T16386.  
 DR WormPep; F47B7.1; CE02743.  
 DR InterPro; IPR000612; UPF0057.  
 DR Pfam; PF01679; UPF0057; 1.  
 DR PROSITE; PS01309; UPF0057; 1.  
 DR Hypothetical protein; Transmembrane.  
 KW TRANSMEM 5 25 POTENTIAL.  
 FT TRANSMEM 37 57 POTENTIAL.  
 SQ SEQUENCE 59 AA; 6840 MW; DCB8905FCE3A660D CRC64;  
 Query Match 57.1%; Score 24; DB 1; Length 59;  
 Best Local Similarity 55.6%; Pred. No. 1.1e+02;  
 Matches 5; Conservative 3; Mismatches 1; Indels 0; Gaps 0;  
 QY 1 NVLMANNII 9  
 DB 30 NMHVAVNII 38  
 RESULT 11  
 ACPH\_UREPA STANDARD; PRT; 77 AA.  
 ID ACPH\_UREPA  
 AC Q9PPY4;  
 DT 28-FEB-2003 (Rel. 41, Created)  
 DT 28-FEB-2003 (Rel. 41, Last sequence update)  
 DT 15-SEP-2003 (Rel. 42, Last annotation update)  
 DE Acyl carrier protein homolog (ACP).  
 DE U0506.  
 GN Ureaplasma parvum (Ureaplasma urealyticum biotype 1).  
 OS Bacteria; Firmicutes; Mollicutes; Mycoplasmataceae; Ureaplasma.  
 OC NCBI\_TaxID=134821;  
 OX [1]  
 RN SEQUENCE FROM N.A.  
 RP STRAIN=Servovir 3; PubMed=11048724;  
 RC MEDLINE=20500219; PubMed=11048724;  
 RX Glass J.I., Lefkowitz E.J., Glass J.S., Heiner C.R., Chen E.Y.,  
 RA Caswell G.H.;  
 RA "The complete sequence of the mucosal pathogen Ureaplasma  
 RT urealyticum.";  
 RT Nature 407:757-762(2000).  
 RL -!- FUNCTION: Carrier of the growing fatty acid chain in fatty acid  
 CC biosynthesis (By similarity).  
 CC -!- PATHWAY: De novo fatty acid biosynthesis.  
 CC -!- PTM: 4'-phosphopantetheine is transferred from CoA to a specific  
 CC serine of the apo-ACP-like protein (potential).  
 CC -!- SIMILARITY: Contains 1 acyl carrier domain.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 CC EMBL; A8002149; AAF30918.1; -  
 DR InterPro; IPR006163; Pf. bind.  
 DR InterPro; IPR006162; Pfam. attach.  
 DR Pfam; PF00550; pp-binding; 1.  
 DR PROSITE; PS00075; ACP DOMAIN; 1.  
 DR PROSITE; PS00012; PHOSPHOPANTHETHEINE; 1.  
 DR Hypothetical protein; Fatty acid biosynthesis; Phosphopantetheine;  
 KW Complete proteome.  
 FT BINDING 36 36 PHOSPHOPANTHETHEINE (BY SIMILARITY).  
 SQ SEQUENCE 77 AA; 8750 MW; 3CDB655F8FA968C6 CRC64;  
 Query Match 57.1%; Score 24; DB 1; Length 77;  
 Best Local Similarity 71.4%; Pred. No. 1.4e+02;  
 Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;  
 QY 3 LMAMNII 9  
 DB 30 NMHVAVNII 38

Db 37 LSMNLI 43  
 RESULT 12  
 FXV3\_HUMAN STANDARD; PRT; 87 AA.  
 ID FXV3\_HUMAN  
 AC Q14802; Q13211;  
 DT 15-JUL-1998 (Rel. 36, Created)  
 DT 15-JUL-1998 (Rel. 36, Last sequence update)  
 DT 15-SEP-2003 (Rel. 42, Last annotation update)  
 DE FXVD domain-containing ion transport regulator 3 precursor (Chloride  
 DE conductance inducer protein Mat-8) (Mammary tumor 8 kDa protein)  
 DE (Phospholemmann-like).  
 DE FXVD3 OR MAT8 OR PLML.  
 GN Homo sapiens (Human)  
 OS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 OX NCBI\_TaxID=9606;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Breast;  
 RX MEDLINE=95138184; PubMed=7836447;  
 RA Morrison B.W., Moorman J.R., Kowdley G.C., Kobayashi Y.M., Jones L.R.,  
 RA Leder P.;  
 RT "Mat-8, a novel phospholemmann-like protein expressed in human breast  
 RT tumors, induces a chloride conductance in Xenopus oocytes.";  
 RL J. Biol. Chem. 270:2176-2182(1995).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Lung carcinoma;  
 RA Lei W., Wu M.;  
 RL Submitted (JUL-1995) to the EMBL/GenBank/DBJ databases.  
 RN [3]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Urinary bladder;  
 RX MEDLINE=22388257; PubMed=12477932;  
 RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,  
 RA Klausner R.D., Collins F.S., Wagner K.H., Schaefer C.F., Bhat N.K.,  
 RA Altschul S.F., Zeeberg B., Buetow K.H., Wang J., Wang J., Hsieh L.,  
 RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hong L.,  
 RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,  
 RA Scapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,  
 RA Brownstein M.J., Ustin T.B., Toshynski S., Carninci P., Prange C.,  
 RA Raha S.S., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,  
 RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,  
 RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,  
 RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,  
 RA Fahey J., Helton E., Kettman M., Madan A., Rodriguez S., Sanchez A.,  
 RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,  
 RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,  
 RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,  
 RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smallos D.E.,  
 RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;  
 RT "Generation and initial analysis of more than 15,000 full-length  
 RT human and mouse cDNA sequences.";  
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).  
 CC -!- FUNCTION: Induces a hyperpolarization-activated chloride current  
 CC when expressed in xenopus oocytes. May be a modulator capable of  
 CC activating endogenous oocyte channels.  
 CC -!- SUBCELLULAR LOCATION: Type I membrane protein (Potential).  
 CC -!- TISSUE SPECIFICITY: Expressed in a subset of human breast tumors.  
 CC -!- MISCELLANEOUS: Marker of a cell type preferentially transformed by  
 CC neu or ras oncoprotein.  
 CC -!- SIMILARITY: Belongs to the FXVD family.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----

```

DR EMBL; X93036; CAA63604.1; -.
DR EMBL; U28249; AAA73922.1; -.
DR EMBL; BC005238; AAH05238.1; -.
DR PIR; A55571; A55571.
DR Genew; HGNC:4027; FXD3.
DR MIM; 604996; -.
DR GO; GO:0005887; C: integral to plasma membrane; TAS.
DR GO; GO:0005254; P: chloride channel activity; TAS.
DR GO; GO:0006821; P: chloride transport; TAS.
DR InterPro; IPR000272; ATP1G1_PLM_MAT8.
DR Pfam; PF02038; ATP1G1_PLM_MAT8; 1.
DR PROSITE; PS01310; FXD; 1.
KW Transmembrane; Signal; Ionic channel; Chloride channel; Chloride;
KW Ion transport.
FT SIGNAL 1 20 POTENTIAL.
FT CHAIN 21 87 FXD DOMAIN-CONTAINING ION TRANSPORT
FT REGULATOR 3.
FT DOMAIN 21 38 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 39 59 POTENTIAL.
FT DOMAIN 60 87 CYTOPLASMIC (POTENTIAL).
FT CONFLICT 36 37 MISSING (IN REF. 2).
FT CONFLICT 58 58 S -> SEWRSSGEQAGRGWGSPPLTTLQSLPTG (IN
FT REF. 2).
FT SEQUENCE 87 AA; 9263 MW; 6D574D668EB32493 CRC64;

Query Match 57.1%; Score 24; DB 1; Length 87;
Best Local Similarity 75.0%; Pred. No. 1.6e+02;
Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 2 VLMMANNII 9
DB 47 VLCAMGII 54

RESULT 13
YFRC_PROVU STANDARD; PRT; 96 AA.
AC P20927;
DT 01-FEB-1991 (Rel. 17, Created)
DT 01-FEB-1991 (Rel. 17, Last sequence update)
DT 01-FEB-1995 (Rel. 31, Last annotation update)
DE FRD operon hypothetical protein C.
OS Proteus vulgaris.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Proteus.
OX NCBI_TaxID=585;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=88004470; PubMed=3308458;
RA Cole S.T.;
RT "Nucleotide sequence and comparative analysis of the frd operon
RT encoding the fumarate reductase of Proteus vulgaris. Extensive
RT sequence divergence of the membrane anchors and absence of an
RT frd-linked ampC cephalosporinase gene.";
RL Eur. J. Biochem. 167:481-488 (1987).
CC -1- SIMILARITY: BELONGS TO THE HUPF/HYPC FAMILY.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; X06151; CAA29511.1; -.
DR PIR; S00119; S00119.
DR InterPro; IPR001109; HupF HupC.
DR Pfam; PF01455; HupF_HypC; 1.
DR PIRSF; PIRSF005618; HupF_HypC; 1.
DR PRINTS; PR00445; HUPFHYP.
DR ProDom; PD003112; HupF_HypC; 1.
DR TIGRFAMs; TIGR00074; hupC_hupF; 1.

DR PROSITE; PS01097; HUPF_HYPC; 1.
KW Hypothetical protein.
SQ SEQUENCE 96 AA; 10300 MW; EC094F1F37956EE3 CRC64;

Query Match 57.1%; Score 24; DB 1; Length 96;
Best Local Similarity 83.3%; Pred. No. 1.8e+02;
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 NVLMAM 6
DB 70 NALMAM 75

RESULT 14
NULM_BALPH STANDARD; PRT; 98 AA.
AC P24976;
DT 01-MAR-1992 (Rel. 21, Created)
DT 01-MAR-1992 (Rel. 21, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE NADH-ubiquinone oxidoreductase chain 4L (EC 1.6.5.3).
GN MTND4L OR ND4L OR NADH4L.
OS Balaenoptera physalus (Finback whale) (Common rorqual).
OG Mitochondrion.
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Cetacea; Mysticeti;
OC Balaenopteridae; Balaenoptera.
OX NCBI_TaxID=9770;
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN=Isolate No. 27 / Anno 1987; TISSUE=Liver;
RX MEDLINE=92139449; PubMed=1779436;
RA Arnason U., Gullberg A., Widegren B.;
RT "The complete nucleotide sequence of the mitochondrial DNA of the fin
RT whale, Balaenoptera physalus.";
RL J. Mol. Evol. 33:556-568 (1991).
CC -1- CATALYTIC ACTIVITY: NADH + ubiquinone = NAD(+) + ubiquinol.
CC -----
DR EMBL; X61145; CAA43447.1; -.
DR PIR; A58851; S24920.
DR InterPro; IPR001133; Oxidored4L.
DR InterPro; IPR003214; Oxidred4L.
DR Pfam; PF00420; Oxidored G2; 1.
DR ProDom; PD000359; Oxidred4L; 1.
DR Oxidoreductase; NAD; Ubiquinone; Mitochondrion.
KW Oxidoreductase; NAD; Ubiquinone; Mitochondrion.
SQ SEQUENCE 98 AA; 10763 MW; 0070D3D6C3AE805F CRC64;

Query Match 57.1%; Score 24; DB 1; Length 98;
Best Local Similarity 50.0%; Pred. No. 1.8e+02;
Matches 4; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 1 NVLMAMNI 8
DB 7 NILMAFSM 14

RESULT 15
NULM_DIDMA STANDARD; PRT; 98 AA.
AC P41307;
DT 01-FEB-1995 (Rel. 31, Created)
DT 01-FEB-1995 (Rel. 31, Last sequence update)
DT 01-NOV-1997 (Rel. 35, Last annotation update)
DE NADH-ubiquinone oxidoreductase chain 4L (EC 1.6.5.3).
GN MTND4L OR ND4L.

```

OS Didelphis marsupialis virginiana (North American opossum).  
 OG Mitochondrion.  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Metatheria; Didelphimorphia; Didelphidae; Didelphis.  
 OX NCBI\_TaxID=9267;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Liver;  
 RX MEDLINE=94333786; PubMed=8056314;  
 RA Janke A., Feldmaier-Fuchs G., Thomas K., von Haeseler A., Paabo S.;  
 RT "The marsupial mitochondrial genome and the evolution of placental  
 mammals";  
 RL Genetics 137:243-256(1994).  
 CC -1- CATALYTIC ACTIVITY: NADH + ubiquinone = NAD(+) + ubiquinol.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See http://www.isb-sib.ch/announce/  
 CC or send an email to license@isb-sib.ch).  
 CC -----  
 CC EMBL; Z29573; CAA82685.1; --  
 DR PIR; S47878; S47878.  
 DR InterPro; IPR001133; Oxidored 4L.  
 DR InterPro; IPR003214; Oxidred4L.  
 DR Pfam; PF00420; Oxidored g2; 1.  
 DR ProDom; PD000359; Oxidred4L; 1.  
 DR Oxidoreductase; NAD; Ubiquinone; Mitochondrion.  
 KW SEQUENCE 98 AA; 10815 MW; F98F602680AC59 CRC64;  
 SQ  
 Query Match 57.1%; Score 24; DB 1; Length 98;  
 Best Local Similarity 62.5%; Pred. No. 1.8e+02;  
 Matches 5; Conservative 2; Mismatches 1; Indels 0; Gaps 0;  
 QY 2 VLIMANNII 9  
 DB . 2 VLILKLI 9  
 RESULT 16  
 ATP8\_CANPA STANDARD; PRT; 48 AA.  
 ID Y666\_CHLTPR  
 AC O84673;  
 DT 01-AUG-1990 (Rel. 15, Created)  
 DT 01-AUG-1990 (Rel. 15, Last sequence update)  
 DT 28-FEB-2003 (Rel. 41, Last annotation update)  
 DE ATP synthase protein 8 (EC 3.6.3.14) (ATPase subunit 8) (A6L).  
 GN ATP8 OR AAP1.  
 OS Candida parapsilosis (Yeast).  
 OG Mitochondrion.  
 OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;  
 OC Saccharomycetales; Mitosporic Saccharomycetales; Candida.  
 OX NCBI\_TaxID=5480;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=CBS 7154 / SPI;  
 RX MEDLINE=90332440; PubMed=2143015;  
 RA Guelin E., Velours J., Guerin M.;  
 RT "Cloning and sequencing of a fragment of the linear mitochondrial DNA  
 RT of the yeast Candida parapsilosis supporting genes encoding subunit 8  
 RT of FO ATP synthase and a putative t-RNA(Pro).";  
 RL Nucleic Acids Res. 18:4267-4267(1990).  
 CC -1- FUNCTION: THIS IS ONE OF THE CHAINS OF THE NONENZYMATIC COMPONENT  
 CC (CF1O) SUBUNIT OF THE MITOCHONDRIAL ATPASE COMPLEX.  
 CC -1- CATALYTIC ACTIVITY: ATP + H(2)O + H(+) (In) = ADP + phosphate +  
 CC H(+) (Out).  
 CC -1- SUBCELLULAR LOCATION: Membrane-bound.  
 CC -1- SIMILARITY: BELONGS TO THE ATPASE PROTEIN 8 FAMILY.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -

CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See http://www.isb-sib.ch/announce/  
 CC or send an email to license@isb-sib.ch).  
 CC -----  
 CC EMBL; X52115; CAA36361.1; --  
 DR PIR; S10465; PWCK8P.  
 KW Hydrogen ion transport; CF(0); Mitochondrion; Transmembrane.  
 FT TRANSMEM 12 32  
 SQ SEQUENCE 48 AA; 5492 MW; 0E7876341CAF0747 CRC64;  
 QY  
 DB

Query Match 54.8%; Score 23; DB 1; Length 48;  
 Best Local Similarity 62.5%; Pred. No. 1.5e+02;  
 Matches 5; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 2 VLIMANNII 9  
 DB 38 LLIARNII 45

## RESULT 17

Y666\_CHLTPR STANDARD; PRT; 83 AA.  
 ID Y666\_CHLTPR  
 AC O84673;  
 DT 16-OCT-2001 (Rel. 40, Created)  
 DT 16-OCT-2001 (Rel. 40, Last sequence update)  
 DT 28-FEB-2003 (Rel. 41, Last annotation update)  
 DE Protein CT666.  
 GN CT666.  
 OS Chlamydia trachomatis.  
 OC Bacteria; Chlamydiae; Chlamydiales; Chlamydiaceae; Chlamydia.  
 OX NCBI\_TaxID=813;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=D/UM-3/Cx;  
 RX MEDLINE=9900809; PubMed=9784136;  
 RA Stephens R.S., Kalman S., Lammel C.J., Fan J., Marathe R., Aravind L.,  
 RA Mitchell W.P., Olinger L., Tatusov R.L., Zhao Q., Koonin E.V.,  
 RA Davis R.W.;  
 RT "Genome sequence of an obligate intracellular pathogen of humans:  
 RT Chlamydia trachomatis";  
 RL Science 282:754-759(1998).  
 CC -1- SIMILARITY: BELONGS TO THE CHLAMYDIAL CPN0710/CT666/TC0037  
 CC FAMILY.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See http://www.isb-sib.ch/announce/  
 CC or send an email to license@isb-sib.ch).  
 CC -----  
 CC EMBL; AE001337; AAC68261.1; --  
 DR PIR; F71485; F71485.  
 DR PHCI-2DPAGE; O84673; --  
 KW Complete proteome.  
 SQ SEQUENCE 83 AA; 9160 MW; FF227A1F3549AAE3 CRC64;  
 QY  
 DB

Query Match 54.8%; Score 23; DB 1; Length 83;  
 Best Local Similarity 57.1%; Pred. No. 2.6e+02;  
 Matches 4; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 NVLMANN 7  
 DB 64 NILTAVN 70

## RESULT 18

Y710\_CHLTPN STANDARD; PRT; 84 AA.  
 ID Y710\_CHLTPN  
 AC Q927J5;

```

DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Protein CPN0710/CP0036/CPJ0710.
GN CPN0710 OR CP0036 OR CPJ0710.
OS Chlamydia pneumoniae (Chlamydophila pneumoniae).
OC Bacteria; Chlamydiae; Chlamydiales; Chlamydiaceae; Chlamydophila.
OX NCBI_TaxID=83558;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CWL029;
RX MEDLINE=99206606; PubMed=10192388;
RA Kalman S., Mitchell W., Marache R., Lammel C., Fan J., Hyman R.W.,
RA Olinger L., Grimwood J., Davis R.W., Stephens R.S.;
RT "Comparative genomes of Chlamydia pneumoniae and C. trachomatis.";
RL Nat. Genet. 21:385-389(1999).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=AR39;
RX MEDLINE=20150255; PubMed=10684935;
RA Read T.D., Brunham R.C., Shen C., Gill S.R., Heidelberg J.F.,
RA White O., Hickey E.K., Peterson J., Utterback T., Berry K., Bass S.,
RA Linher K., Weidman J., Khouri H., Craven B., Bowman C., Dodson R.,
RA Winn M., Nelson W., DeBoy R., Kolonay J., McClarty G., Salzberg S.L.,
RA Eisen J., Fraser C.M.;
RT "Genome sequences of Chlamydia trachomatis MoPn and Chlamydia
RT pneumoniae AR39.";
RL Nucleic Acids Res. 28:1397-1406(2000).
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=J138;
RX MEDLINE=20330349; PubMed=10871362;
RA Shirai M., Hirakawa H., Kimoto M., Tabuchi M., Kishi F., Ouchi K.,
RA Shiba T., Ishii K., Hattori K., Kuwara S., Nakazawa T.;
RT "Comparison of whole genome sequences of Chlamydia pneumoniae J138
RT from Japan and CWL029 from USA.";
RL Nucleic Acids Res. 28:2311-2314(2000).
RN [4]
RP SEQUENCE FROM N.A.
RC STRAIN=J138;
RA Hirakawa H., Shirai M., Kuwara S.;
RT "Genomic sequence comparison of two unrelated isolates of Chlamydia
RT pneumoniae from Japan and U.S.";
RL Submitted (JAN-2000) to the EMBL/GenBank/DBJ databases.
CC -1- SIMILARITY: BELONGS TO THE CHLAMYDIAL CPN0710/CPJ0710/TC0037
CC FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; AE001652; AAD18849.1; -
DR EMBL; AE002167; AAF37931.1; -
DR EMBL; AP002547; BAA98917.1; -
DR EMBL; AB035952; BAA88660.1; -
DR FIR; C86579; C86579.
DR FIR; F72046; F72046.
DR PHCI-2DPAGE; Q92705; -.
DR TIGR; CP0036; -.
KW Complete proteome.
SQ SEQUENCE 84 AA; 9348 MW; 50A9CC271A5CB693 CRC64;
Query Match 54.8%; Score 23; DB 1; Length 84;
Best Local Similarity 57.1%; Pred. No. 2.6e+02;
Matches 4; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

OY 1 NVLMANN 7
Db 65 NILTAVN 71

```

```

* RESULT 19
ID _RPOH_THEAC STANDARD; PRT; 87 AA.
AC Q03588;
DT 01-JUN-1994 (Rel. 29, Created)
DT 01-JUN-1994 (Rel. 29, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE DNA-directed RNA polymerase subunit H (EC 2.7.7.6).
GN RPOH OR TA0389
OS Thermoplasma acidophilum.
OC Archaea; Euryarchaeota; Thermoplasmata; Thermoplasmatales;
OC Thermoplasmataceae; Thermoplasma.
OX NCBI_TaxID=2303;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=DSM 1728;
RX MEDLINE=93027268; PubMed=1408839;
RA Klenk H.-P., Renner O., Schwass V., Zillig W.;
RT "Nucleotide sequence of the genes encoding the subunits H, B, A' and
RT A'' of the DNA-dependent RNA polymerase and the initiator tRNA from
RT Thermoplasma acidophilum.";
RL Nucleic Acids Res. 20:5226-5226(1992).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=DSM 1728;
RX MEDLINE=20479972; PubMed=11029001;
RA Ruepp A., Graml W., Santos-Martinez M.-L., Koretke K.K., Volker C.,
RA Mewes H.-W., Frishman D., Stocker S., Lupas A.N., Baumeister W.;
RT "The genome sequence of the thermoacidophilic scavenger Thermoplasma
RT acidophilum.";
RL Nature 407:508-513(2000).
CC -1- FUNCTION: DNA-DEPENDENT RNA POLYMERASE CATALYZES THE TRANSCRIPTION
CC OF DNA INTO RNA USING THE FOUR RIBONUCLEOSIDE TRIPHOSPHATES AS
CC SUBSTRATES.
CC -1- CATALYTIC ACTIVITY: N nucleoside triphosphate = N diphosphate +
CC {RNA} (N).
CC -1- SIMILARITY: BELONGS TO THE ARCHAEABACTERIA RPOH / EUKARYOTIC RPB5
CC RNA POLYMERASE SUBUNIT FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; X68198; CAA48279.1; -
DR EMBL; AL445064; CAC11533.1; ALT_INIT.
DR FIR; S26721; S26721.
DR HSSP; O27122; IEIK.
DR HAMAP; MF_00025; -.
DR InterPro; IPR000783; RNA_pol_H_23kd.
DR Pfam; PF01191; RNA_pol_Rpb5_C; 1.
DR PRODOM; PD005155; RNA_pol_H_23kd; 1.
DR PROSITE; PS01110; RNA_POL_H_23KD; 1.
KW Transferase; Transcription; DNA-directed RNA polymerase;
KW Complete proteome.
SQ SEQUENCE 87 AA; 9927 MW; 9C5B1F39ABAC6F4E CRC64;
Query Match 54.8%; Score 23; DB 1; Length 87;
Best Local Similarity 50.0%; Pred. No. 2.7e+02;
Matches 4; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

OY 1 NVLMANNI 8
Db 25 NILKELNI 32

```

```

RESULT 20
RL31_HALNI

```

```

ID  RL31 HALN1          STANDARD;          PRT;          91 AA.
AC  Q9H9N0;
DT  16-OCT-2001 (Rel. 40, Created)
DT  16-OCT-2001 (Rel. 40, Last sequence update)
DT  28-FEB-2003 (Rel. 41, Last annotation update)
DE  50S ribosomal protein L31e.
GN  RPL31E OR VNG2467G.
OS  Halobacterium sp. (strain NRC-1 / ATCC 700922 / JCM 11081).
OC  Archaea; Euryarchaeota; Halobacteria; Halobacteriales;
OC  Halobacteriaceae; Halobacterium.
OX  NCBI_TaxID=64091;
RN  [1]
RP  SEQUENCE FROM N.A.
RX  MEDLINE=20504483; PubMed=11016950;
RA  Ng W.V., Kennedy S.P., Mahaisas G.G., Berquist B., Pan M.,
RA  Shukla H.D., Laaky S.R., Baliga N.S., Thorsson V., Sbrogna J.,
RA  Swartzell S., Weir D.R., Hall J., Dahl T.A., Welter R., Goo Y.A.,
RA  Leithauser B., Keller K., Cruz R., Danson M.J., Hough D.W.,
RA  Maddocks D.G., Jablonski P.E., Krebs M.P., Angevine C.M., Dale H.,
RA  Isenbarger T.A., Peck R.F., Pohlschroder M., Spudich J.L., Jung K.-H.,
RA  Alam M., Freitas T., Hou S., Daniels C.J., Dennis P.P., Omer A.D.,
RA  Ehardt H., Lowe T.M., Liang P., Riley M., Hood L., DasSarma S.;
RA  "Genome sequence of Halobacterium species NRC-1.";
RT  Proc. Natl. Acad. Sci. U.S.A. 97:12176-12181(2000).
RL  -!- SIMILARITY: BELONGS TO THE L31E FAMILY OF RIBOSOMAL PROTEINS.
CC  -----
CC  This SWISS-PROT entry is copyright. It is produced through a collaboration
CC  between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC  the European Bioinformatics Institute. There are no restrictions on its
CC  use by non-profit institutions as long as its content is in no way
CC  modified and this statement is not removed. Usage by and for commercial
CC  entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC  or send an email to license@isb-sib.ch).
CC  -----
DR  EMBL; AE005125; AAG20541.1; -.
DR  PIR; A84397; A84397.
DR  HAMAP; MF 00410; -.
DR  InterPro; IPR000054; Ribosomal_L31e.
DR  Pfam; PF01198; Ribosomal_L31e; 1.
DR  PROSITE; PS01144; RIBOSOMAL_L31e; 1.
KW  Ribosomal protein; Complete proteome.
FT  INIT MET 0 BY SIMILARITY.
SQ  SEQUENCE 91 AA; 10092 MW; 25003B509C7F99D5 CRC64;

Query Match 54.8%; Score 23; DB 1; Length 91;
Best Local Similarity 100.0%; Pred. No. 2.8e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY  5 AMNII 9
DB  29 AMNII 33

RESULT 21
VAPD_VAPD_HAEIN STANDARD; PRT; 91 AA.
AC  P71351;
DT  15-JUL-1999 (Rel. 38, Created)
DT  15-JUL-1999 (Rel. 38, Last sequence update)
DT  16-OCT-2001 (Rel. 40, Last annotation update)
DE  Virulence-associated protein D.
GN  VAPD OR HI0450.
OS  Haemophilus influenzae.
OC  Bacteria; Proteobacteria; Gammaproteobacteria; Pasteurellales;
OC  Pasteurellaceae; Haemophilus.
OX  NCBI_TaxID=727;
RN  [1]
RP  SEQUENCE FROM N.A.
RX  STRAIN=rd / KW20 / ATCC 51907;
RX  MEDLINE=9530630; PubMed=7542800;
RA  Fleischmann R.D., Adams M.D., White O., Clayton R.A., Kirkness E.F.,
RA  Kerlavage A.R., Bult C.J., Tomb J.-P., Dougherty B.A., Merrick J.M.,
RA  McKenney K., Sutton G., Fitzhugh W., Fields C.A., Gocayne J.D.,

```

```

RA  Scott J.D., Shirley R., Liu L.-I., Glodek A., Kelley J.M.,
RA  Weidman J.F., Phillips C.A., Spriggs T., Hedblom E., Cotton M.D.,
RA  Uterback T.R., Hanna M.C., Nguyen D.T., Saudek D.M., Brandon R.C.,
RA  Fine L.D., Fritchman J.L., Fuhrmann J.L., Geoghagen N.S.M.,
RA  Gnehm C.L., McDonald L.A., Small K.V., Fraser C.M., Smith H.O.,
RA  Venter J.C.;
RT  "Whole-genome random sequencing and assembly of Haemophilus influenzae
RL  Rd.";
RL  Science 269:496-512(1995).
RN  [2]
RP  SEQUENCE OF 76-80.
RX  MEDLINE=20137488; PubMed=10675023;
RX  Langen H., Takacs B., Evers S., Berndt P., Lehm H.W., Wipf B.,
RA  Gray C., Fountoulakis M.;
RT  "Two-dimensional map of the proteome of Haemophilus influenzae.";
RL  Electrophoresis 21:411-429(2000).
CC  -!- SIMILARITY: BELONGS TO THE VAPD FAMILY.
CC  -----
CC  This SWISS-PROT entry is copyright. It is produced through a collaboration
CC  between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC  the European Bioinformatics Institute. There are no restrictions on its
CC  use by non-profit institutions as long as its content is in no way
CC  modified and this statement is not removed. Usage by and for commercial
CC  entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC  or send an email to license@isb-sib.ch).
CC  -----
DR  EMBL; U32728; AAC22108.1; -.
DR  PIR; C64069; C64069.
DR  TIGR; HI0450; -.
DR  Pfam; PF04605; VapD_N; 1.
KW  Virulence; Complete proteome.
SQ  SEQUENCE 91 AA; 10543 MW; 70B23CDE28E083E7 CRC64;

Query Match 54.8%; Score 23; DB 1; Length 91;
Best Local Similarity 44.4%; Pred. No. 2.8e+02;
Matches 4; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY  1 NVLMAMNII 9
DB  53 NLFQAMNII 61

RESULT 22
Y736_ARCFU STANDARD; PRT; 95 AA.
AC  O29522;
DT  16-OCT-2001 (Rel. 40, Created)
DT  16-OCT-2001 (Rel. 40, Last sequence update)
DT  16-OCT-2001 (Rel. 40, Last annotation update)
DE  Hypothetical protein AF0736.
GN  AF0736.
OS  Archaeoglobus fulgidus.
OC  Archaea; Euryarchaeota; Archaeoglobi; Archaeoglobales;
OC  Archaeoglobaceae; Archaeoglobus.
OX  NCBI_TaxID=2234;
RN  [1]
RP  SEQUENCE FROM N.A.
RX  STRAIN=VC-16 / DSM 4304 / ATCC 49558;
RX  MEDLINE=98049343; PubMed=9389475;
RA  Klenk H.-P., Clayton R.A., Tomb J.-F., White O., Nelson K.E.,
RA  Ketchum K.A., Dodson R.J., Gwinn M., Hickey E.K., Peterson J.D.,
RA  Richardson D.L., Kerlavage A.R., Graham D.E., Kyrpides N.C.,
RA  Fleischmann R.D., Quackenbush J., Lee N.H., Sutton G.G., Gill S.,
RA  Kirkness E.F., Dougherty B.A., McKenney K., Adams M.D., Loftus B.,
RA  Peterson S., Reich C.I., McNeil L.K., Badger J.H., Glodek A., Zhou L.,
RA  Overbeek R., Gocayne J.D., Weidman J.F., McDonald L., Uterback T.,
RA  Cotton M.D., Spriggs T., Artach P., Kaine B.P., Sykes S.M.,
RA  Sadow P.W., D'Andrea K.P., Bowman C., Fujii C., Garland S.A.,
RA  Mason T.M., Olsen G.J., Fraser C.M., Smith H.O., Woese C.R.,
RA  Venter J.C.;
RT  "The complete genome sequence of the hyperthermophilic, sulphate-
RT  reducing archaeon Archaeoglobus fulgidus.";
RL  Nature 390:364-370(1997).

```

```
CC -1- SUBCELLULAR LOCATION: Integral membrane protein (Potential).
CC -1- SIMILARITY: BELONGS TO THE UPF0132 FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; AF001054; AAB90510.1; -.
CC PIR; H69341; H69341.
CC TIGR; AF0736; -.
CC InterPro; IPR005348; UPF0132.
CC Pfam; PF03675; UPF0132; 1.
CC ProDom; PD015391; UPF0132; 1.
CC Hypothetical protein; Transmembrane; Complete proteome.
KW TRANSMEM 2 22 POTENTIAL.
FT TRANSMEM 32 52
FT TRANSMEM 55 75 POTENTIAL.
SQ SEQUENCE 95 AA; 10475 MW; 310B66160F26858A CRC64;

Query Match 54.8%; Score 23; DB 1; Length 95;
Best Local Similarity 62.5%; Pred. No. 3e+02;
Matches 5; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 2 VLMAMNII 9
DB 55 VVMAFSII 62
|::|||

RESULT 23
NULM_BOVIN
ID NULM_BOVIN STANDARD; PRT; 98 AA.
AC P03902; Q8SFX7;
DT 21-JUL-1986 (Rel. 01, Created)
DT 21-JUL-1986 (Rel. 01, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE NADH-ubiquinone oxidoreductase chain 4L (EC 1.6.5.3).
GN MTND4L OR ND4L.
OS Bos taurus (Bovine).
OG Mitochondrion.
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovoidea;
OC Bovidae; Bovinae; Bos.
OX NCBI_TaxID=9913;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Heart;
RX MEDLINE=83010260; PubMed=7120390;
RA Anderson S., de Bruijn M.H.L., Coulson A.R., Eperon I.C., Sanger F.,
RA Young I.G.;
RT "Complete sequence of bovine mitochondrial DNA. Conserved features of
RT the mammalian mitochondrial genome.";
RL J. Mol. Biol. 156:683-717(1982).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=65, 66, D, and F;
RA Wettstein P.J.;
RA "Bos taurus mitochondrial protein coding regions.";
RL Submitted (MAR-2002) to the EMBL/GenBank/DBJ databases.
CC -1- CATALYTIC ACTIVITY: NADH + ubiquinone = NAD(+) + ubiquinol.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; V00654; CAA24004.1; -.
CC
```

```
DR EMBL; J01394; AAB59276.1; -.
DR EMBL; AF490528; AAM08325.1; -.
DR EMBL; AF490529; AAM08338.1; -.
DR EMBL; AF493541; AAM12797.1; -.
DR EMBL; AF493542; AAM12810.1; -.
DR PIR; A00429; QX8041.
DR InterPro; IPR001133; Oxidored4L.
DR InterPro; IPR003214; Oxidred4L.
DR Pfam; PF00420; Oxidored_q2; 1.
DR ProDom; PD000359; Oxidred4L; 1.
KW Oxidoreductase; NAD; Ubiquinone; Mitochondrion.
FT VARIANT 65 65 V -> A (IN STRAIN 66).
SQ SEQUENCE 98 AA; 10797 MW; 4DB48B7DA59C1881 CRC64;

Query Match 54.8%; Score 23; DB 1; Length 98;
Best Local Similarity 37.5%; Pred. No. 3.1e+02;
Matches 3; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 1 NVLMAMNI 8
DB 7 NIMMAFTV 14
|::|||

RESULT 24
NULM_SHEEP
ID NULM_SHEEP STANDARD; PRT; 98 AA.
AC O78754;
DT 15-DEC-1998 (Rel. 37, Created)
DT 15-DEC-1998 (Rel. 37, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE NADH-ubiquinone oxidoreductase chain 4L (EC 1.6.5.3).
GN MTND4L OR ND4L.
OS Ovis aries (Sheep).
OG Mitochondrion.
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovoidea;
OC Bovidae; Caprinae; Ovis.
OX NCBI_TaxID=9940;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Merinolandschaf; TISSUE=Liver;
RX MEDLINE=98440761; PubMed=9767689;
RA Hiendleder S., Lewalski H., Wassmuth R., Janke A.;
RT "The complete mitochondrial DNA sequence of the domestic sheep (Ovis
RT aries) and comparison with the other major ovine haplotype.";
RL J. Mol. Evol. 47:441-448(1998).
CC -1- CATALYTIC ACTIVITY: NADH + ubiquinone = NAD(+) + ubiquinol.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; AF010406; AAD10103.1; -.
DR PIR; T11058; T11058.
DR InterPro; IPR001133; Oxidored4L.
DR InterPro; IPR003214; Oxidred4L.
DR Pfam; PF00420; Oxidored_q2; 1.
DR ProDom; PD000359; Oxidred4L; 1.
KW Oxidoreductase; NAD; Ubiquinone; Mitochondrion.
SQ SEQUENCE 98 AA; 10837 MW; ECOC943C752F8691 CRC64;

Query Match 54.8%; Score 23; DB 1; Length 98;
Best Local Similarity 37.5%; Pred. No. 3.1e+02;
Matches 3; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 1 NVLMAMNI 8
DB 7 NIMMAFTV 14
|::|||
```



```

[1]
RN  SEQUENCE FROM N.A.
RP  STRAIN=JAL-1 / DSM 2661 / ATCC 43067;
RX  MEDLINE=9637999; PubMed=8688087; Zhou L., Fleischmann R.D., Gocayne J.D.,
RA  Sutton G.G., White O., Olsen G.J., FitzGerald L.M., Clayton R.A., Adams M.D., Reich C.I.,
RA  Kerlavage A.R., Dougherty B.A., Tomb J.-F., Adams M.D., Merrick J.M., Glodek A.,
RA  Overbeek R., Kirkness E.F., Weinstock K.G., Weidman J.P., Fuhrmann J.L., Nguyen D.,
RA  Scott J.L., Geoghegan N.S.M., Peterson J.D., Sadow P.W., Hanna M.C.,
RA  Uterback T.R., Kelley J.M., Hurst M.A., Kaine B.P., Borodovsky M.,
RA  Cotton M.D., Roberts K.M., Smith H.O., Woese C.R., Venter J.C.,
RA  Klenk H.-P., Fraser C.M., Smith H.O., Woese C.R., Venter J.C.,
RT  "Complete genome sequence of the methanogenic archaeon, Methanococcus
RL  jannaschii";
CC  Science 273:1058-1073(1996).
CC  -1- SUBCELLULAR LOCATION: Integral membrane protein (Potential).
CC  This SWISS-PROT entry is copyright. It is produced through a collaboration
CC  between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC  the European Bioinformatics Institute. There are no restrictions on its
CC  use by non-profit institutions as long as its content is in no way
CC  modified and this statement is not removed. Usage by and for commercial
CC  entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC  or send an email to license@isb-sib.ch).
CC  -----
DR  EMBL; U67592; AAB99536.1; -.
DR  PIR; C64489; C64489.
DR  TIGR; M1516; -.
DR  Pfam; PF0154; Tm2; 1.
KW  Hypothetical protein; Transmembrane; Complete proteome.
FT  TRANSMEM 34 54
FT  TRANSMEM 56 76
FT  TRANSMEM 99 AA; 11483 MW; B837A92FA94002FB CRC64;
SQ  SEQUENCE 99 AA; 11483 MW; B837A92FA94002FB CRC64;
Query Match 54.8%; Score 23; DB 1; Length 99;
Best Local Similarity 71.4%; Pred. No. 3.1e+02;
Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
QY 3 LMANNII 9
DB 1 WMANNI 7
RESULT 27
ID  ADHR_DROYA STANDARD; PRT; 32 AA.
AC  P28487;
DT  01-DEC-1992 (Rel. 24, Created)
DT  01-DEC-1992 (Rel. 24, Last sequence update)
DT  01-OCT-1996 (Rel. 34, Last annotation update)
DE  Alcohol dehydrogenase related 31 kDa protein (Fragment).
GN  ADHR.
OS  Drosophila yakuba (Fruit fly).
OC  Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC  Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC  Ephydroidea; Drosophilidae; Drosophila.
OX  NCBI_TaxID=7245;
RN  SEQUENCE FROM N.A.
RP  Ashburner M.;
RL  Submitted (JUL-1990) to the EMBL/GenBank/DBJ databases.
CC  -1- SIMILARITY: Belongs to the short-chain dehydrogenases/reductases
CC  (SDR) family.
CC  This SWISS-PROT entry is copyright. It is produced through a collaboration
CC  between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC  the European Bioinformatics Institute. There are no restrictions on its
CC  use by non-profit institutions as long as its content is in no way
CC  modified and this statement is not removed. Usage by and for commercial
CC  entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC  or send an email to license@isb-sib.ch).
CC  -----
DR  EMBL; X54120; CAA38064.1; -.

```

```

RESULT 25
RL3E_METH  STANDARD; PRT; 98 AA.
ID  Q27127;
AC  Q27127;
DT  15-JUL-1998 (Rel. 36, Created)
DT  15-JUL-1998 (Rel. 36, Last sequence update)
DT  28-FEB-2003 (Rel. 41, Last annotation update)
DE  50S ribosomal protein L30e.
GN  RPL30E OR MTH1053.
OS  Methanobacterium thermoautotrophicum.
OC  Archaea; Euryarchaeota; Methanobacteria; Methanobacteriales;
OC  Methanobacteriaceae; Methanothermobacter.
OX  NCBI_TaxID=187420;
RN  SEQUENCE FROM N.A.
RP  STRAIN=Delta H;
RX  MEDLINE=98037514; PubMed=9371463;
RA  Smith D.R., Doucette-Stamm L.A., Deloughery C., Lee H.-M., Dubois J.,
RA  Aldredge T., Bashizadeh R., Blakely D., Cook R., Gilbert K.,
RA  Harrison D., Hoang L., Keagle P., Lumm W., Pothier B., Qiu D.,
RA  Spadafora R., Vicare R., Wang Y., Wierzbowski J., Gibson R.,
RA  Jiwan N., Caruso A., Bush D., Safer H., Patwell D., Prabhakar S.,
RA  McDougall S., Shimer G., Goyal A., Pietrowski S., Church G.M.,
RA  Daniels C.J., Mao J.-I., Rice P., Noelling J., Reeve J.N.;
RA  "Complete genome sequence of Methanobacterium thermoautotrophicum
RT  deltaH: functional analysis and comparative genomics.";
RL  J. Bacteriol. 179:7135-7155(1997).
CC  -1- SIMILARITY: BELONGS TO THE L30E FAMILY OF RIBOSOMAL PROTEINS.
CC  This SWISS-PROT entry is copyright. It is produced through a collaboration
CC  between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC  the European Bioinformatics Institute. There are no restrictions on its
CC  use by non-profit institutions as long as its content is in no way
CC  modified and this statement is not removed. Usage by and for commercial
CC  entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC  or send an email to license@isb-sib.ch).
CC  -----
DR  EMBL; AE000876; AAB85544.1; -.
DR  PIR; A69007; A69007.
DR  HSSP; P14120; 1CN9.
DR  HAMAP; MF 00481; -.
DR  InterPro; IPR000231; Ribosomal L30e.
DR  InterPro; IPR004038; Ribosomal L7A.
DR  Pfam; PF01248; Ribosomal L7Ae; 1.
DR  ProDom; PD004495; Ribosomal L30e; 1.
DR  PROSITE; PS00709; RIBOSOMAL L30E_1; 1.
DR  PROSITE; PS00993; RIBOSOMAL L30E_2; 1.
KW  Ribosomal protein; Complete proteome.
SQ  SEQUENCE 98 AA; 10506 MW; A098F7AB1B29D2F9 CRC64;
Query Match 54.8%; Score 23; DB 1; Length 98;
Best Local Similarity 71.4%; Pred. No. 3.1e+02;
Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
QY 2 VLMANNI 8
DB 35 VWMANNI 41
RESULT 26
YF16_METJA  STANDARD; PRT; 99 AA.
ID  Q58911;
AC  Q58911;
DT  16-OCT-2001 (Rel. 40, Created)
DT  16-OCT-2001 (Rel. 40, Last sequence update)
DT  16-OCT-2001 (Rel. 40, Last annotation update)
DE  Hypothetical protein M1516.
GN  M1516.
OS  Methanococcus jannaschii.
OC  Archaea; Euryarchaeota; Methanococci; Methanococcales;
OC  Methanococcaceae; Methanococcaceae; Methanococcus.
OX  NCBI_TaxID=2190;

```



```

DR PIR; S20719; S20719.
DR FlyBase; FBgn0013164; Dyak\Adhr.
DR InterPro; IPR002198; ADH_short.
DR PROSITE; PS00061; ADH_SHORT; PARTIAL.
KW Oxidoreductase.
FT NP_BIND 11 >32 NAD OR NADP (BY SIMILARITY).
FT NON_TER 32
SQ SEQUENCE 32 AA; 3412 MW; 9302949AFCBE7D4F CRC64;

Query Match 52.4%; Score 22; DB 1; Length 32;
Best Local Similarity 71.1%; Pred. No. 1.6e+02;
Matches 5; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 2 VLMAMNI 8
DB 25 VLMTKNI 31

RESULT 28
YJZJ_ECOLI STANDARD; PRT; 78 AA.
AC P55914;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DE Hypothetical protein yjz.
GN YJZJ OR B4367.1.
OS Escherichia coli.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Escherichia.
OX NCBI_TaxID=562;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=K12 / MG1655;
RX MEDLINE=95334362; PubMed=7610040;
RA Burland V.D., Plunkett G. III, Sofia H.J., Daniels D.L.,
RA Blattner F.R.;
RT "Analysis of the Escherichia coli genome VI: DNA sequence of the
RT region from 92.8 through 100 minutes.";
RL Nucleic Acids Res. 23:2105-2119(1995).
RN [2]
RP IDENTIFICATION.
RA Rudd K.E.;
RL Unpublished observations (OCT-1996).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; U14003; -; NOT ANNOTATED_CDS.
DR EMBL; AE00507; -; NOT ANNOTATED_CDS.
DR EcGene; EG13348; yjz.
KW Hypothetical protein; Complete proteome.
SQ SEQUENCE 78 AA; 8697 MW; B466F1E01231D334 CRC64;

Query Match 52.4%; Score 22; DB 1; Length 78;
Best Local Similarity 57.1%; Pred. No. 4e+02;
Matches 4; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 2 VLMAMNI 8
DB 69 MLMIMNL 75

RESULT 29
C2_OXYNO STANDARD; PRT; 82 AA.
ID_C2_OXYNO
AC P05526;
DT 01-NOV-1988 (Rel. 09, Created)
DT 01-NOV-1988 (Rel. 09, Last sequence update)
DT 01-NOV-1988 (Rel. 09, Last annotation update)
DE C2 protein.
GN C2.
OS Oxytricha nova.
OC Eukaryota; Alveolata; Ciliophora; Spirotrichea; Stichotrichia;
OC Stichotrichida; Oxytrichidae; Oxytricha.
OX NCBI_TaxID=5945;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=84156523; PubMed=6323015;
RA Klobutcher L.A., Jahn C.L., Prescott D.M.;
RT "Internal sequences are eliminated from genes during macronuclear
RT development in the ciliated protozoan Oxytricha nova.";
RL Cell 36:1045-1055(1984).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; K02624; AAA64886.1; -
DR EMBL; K02625; AAA64887.1; -
DR EMBL; K02628; AAA64888.1; -
DR InterPro; IPR004217; Znf_Tim10/BDP.
DR Pfam; PF02953; zf-Tim10_BDP; 1
SQ SEQUENCE 82 AA; 9028 MW; 297AFFBC17079438 CRC64;

Query Match 52.4%; Score 22; DB 1; Length 82;
Best Local Similarity 55.6%; Pred. No. 4.3e+02;
Matches 5; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 1 NVLMAMNII 9
DB 61 SAFMAMNDI 69

RESULT 30
S61B_ARATH STANDARD; PRT; 82 AA.
ID S61B_ARATH
AC P38389;
DT 01-OCT-1994 (Rel. 30, Created)
DT 01-OCT-1994 (Rel. 30, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Protein transport protein SEC61 beta subunit.
GN AT2G45070 OR T14P1.12.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eurosids II; Brassicales; Brassicaceae; Arabidopsis.
OX NCBI_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=94150683; PubMed=8107851;
RA Hartmann E., Sommer T., Prehn S., Goerlich D., Jentsch S.,
RA Rapoport T.A.;
RT "Evolutionary conservation of components of the protein translocation
RT complex.";
RL Nature 367:654-657(1994).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=cv. Columbia;
RX MEDLINE=20083487; PubMed=10617197;
RA Lin X., Kaul S., Rounsley S.D., Shea T.P., Benito M.-I., Town C.D.,
RA Fujii C.Y., Mason T.M., Bowman C.L., Barnstead M.E., Feldblum T.V.,
RA Buell C.R., Ketchum K.A., Lee J.J., Ronning C.M., Koo H.L.,
RA Moffat K.S., Cronin L.A., Shen M., Pai G., Van Aken S., Umayam L.,
RA Tallon L.J., Gill J.E., Adams M.D., Carrera A.J., Creasy T.H.,
RA Goodman H.M., Somerville C.R., Copenhaver G.P., Preuss D.,
RA Nierman W.C., White O., Eisen J.A., Salzberg S.L., Fraser C.M.,

```

modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

EMBL; D90900; BAA16719.1; -  
PIR; S74567; S74567.  
HMAP; MF 00474; -; 1.  
InterPro: IPR000549; PSI psaG/K.  
Pfam; PF01241; PSI\_PSAK; 1.  
PROSITE; PS01026; PHOTOSYSTEM\_I\_PSAK; 1.  
Photosystem I; Photosynthesis; Transmembrane; Thylakoid;  
Complete proteome.  
PROPEP 1 ? POTENTIAL.  
CHAIN ? 86 PHOTOSYSTEM I REACTION CENTER SUBUNIT.  
PSAK 1.  
TRANSMEM 14 34 POTENTIAL.  
TRANSMEM 61 81 POTENTIAL.  
SQ SEQUENCE 86 AA; 8644 MW; 5F0BEAD1B9E99E3B CRC64;

Query Match 52.4%; Score 22; DB 1; Length 86;  
Best Local Similarity 50.0%; Pred. No. 4.5e+02;  
Matches 4; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 2 VLMAMNII 9  
DB 23 VMIACNII 30

RESULT 32  
Y069\_AERPE  
ID Y069\_AERPE STANDARD; PRT; 91 AA.  
AC Q9Y9F0;  
DT 16-OCT-2001 (Rel. 40, Created)  
DT 16-OCT-2001 (Rel. 40, Last sequence update)  
DT 15-SEP-2003 (Rel. 42, Last annotation update)  
DE Hypothetical UPF0147 protein APES069.  
GN APES069.  
OS Aeropyrum pernix.  
OC Archaea; Crenarchaeota; Thermoprotei; Desulfurococcales;  
OC Desulfurococceae; Aeropyrum.  
OX NCBI\_TaxID=56636;  
RN [1]  
RC SEQUENCE FROM N.A.  
RP STRAIN=K1;  
RX MEDLINE=99310339; PubMed=10382966;  
RA Kawarabayashi Y., Hino Y., Horikawa H., Yamazaki S., Haikawa Y.,  
Jin-no K., Takahashi M., Sekine M., Baba S.-I., Ankai A., Kosugi H.,  
Hosoyama A., Fukui S., Nagai Y., Nishijima K., Nakazawa H.,  
Takamiya M., Masuda S., Funahashi T., Tanaka T., Kudoh Y.,  
Yamazaki J., Kishida N., Oguchi A., Aoki K.-I., Kubota K.,  
Nakamura Y., Nomura N., Sako Y., Kikuchi H.;  
RT "Complete genome sequence of an aerobic hyper-thermophilic  
Crenarchaeon, Aeropyrum pernix K1."  
RL DNA Res. 6:83-101(1999).  
CC -1- SIMILARITY: Belongs to the UPF0147 family.  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
between the Swiss Institute of Bioinformatics and the EMBL outstation -  
the European Bioinformatics Institute. There are no restrictions on its  
use by non-profit institutions as long as its content is in no way  
modified and this statement is not removed. Usage by and for commercial  
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

EMBL; AP000064; BAA81350.1; -  
PIR; F72461; F72461.  
HMAP; MF 00342; -; 1.  
InterPro: IPR005354; UPF0147.  
Pfam; PF03685; UPF0147; 1.  
ProDom; PD029943; UPF0147; 1.  
HYPOTHETICAL protein; Complete proteome.  
SQ SEQUENCE 91 AA; 10080 MW; B6DE0AF50AFD94A CRC64;

QY 1 NVLMAMNII 8  
DB 56 NVVLIMSI 63

RESULT 31  
PAK1\_SYNY3  
ID PAK1\_SYNY3 STANDARD; PRT; 86 AA.  
AC P727L2;  
DT 01-NOV-1997 (Rel. 35, Created)  
DT 01-NOV-1997 (Rel. 35, Last sequence update)  
DT 28-FEB-2003 (Rel. 41, Last annotation update)  
DE Photosystem I reaction center subunit peak I precursor (Photosystem I  
subunit X 1).  
GN PSAXI OR PSAX OR SSR0390.  
OS Synechocystis sp. (strain PCC 6803).  
OC Bacteria; Cyanobacteria; Chroococcales; Synechocystis.  
OX NCBI\_TaxID=1148;  
RN [1]  
RC SEQUENCE FROM N.A.  
RX MEDLINE=97061201; PubMed=8905231;  
RA Kaneko T., Sato S., Kotani H., Tanaka A., Asamizu E., Nakamura Y.,  
Miyajima N., Hiroseawa M., Sugliura M., Sasamoto S., Kimura T.,  
Hosouchi T., Matsuno A., Muraki A., Nakazaki N., Naruo K.,  
Okumura S., Shimpo S., Takeuchi C., Wada T., Watanabe A.,  
Okamura M., Yasuda M., Tabata S.;  
RT "Sequence analysis of the genome of the unicellular cyanobacterium  
Synechocystis sp. strain PCC6803. II. Sequence determination of the  
entire genome and assignment of potential protein-coding regions."  
RL DNA Res. 3:109-136(1996).  
CC -1- FUNCTION: THE EXACT FUNCTION OF PSAX IN PHOTOSYSTEM I (PSI) IS  
NOT YET KNOWN.  
CC -1- SUBCELLULAR LOCATION: Integral membrane protein. Cellular  
thylakoid membrane (Probable).  
CC -1- SIMILARITY: Belongs to the psaG / psak family.  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
between the Swiss Institute of Bioinformatics and the EMBL outstation -  
the European Bioinformatics Institute. There are no restrictions on its  
use by non-profit institutions as long as its content is in no way

Venter J.C.;  
RT "Sequence and analysis of chromosome 2 of the plant Arabidopsis  
thaliana."  
RL Nature 402:761-768(1999).  
CC -1- FUNCTION: NECESSARY FOR PROTEIN TRANSLOCATION IN THE ENDOPLASMIC  
RETICULUM.  
CC -1- SUBUNIT: HETEROTRIMERIC COMPLEX COMPOSED OF SEC61-ALPHA,  
SEC61-BETA AND SEC61-GAMMA.  
CC -1- SUBCELLULAR LOCATION: Tail-anchored membrane protein.  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
between the Swiss Institute of Bioinformatics and the EMBL outstation -  
the European Bioinformatics Institute. There are no restrictions on its  
use by non-profit institutions as long as its content is in no way  
modified and this statement is not removed. Usage by and for commercial  
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

EMBL; Z26753; CAA81412.1; -  
EMBL; AC007659; AAD32829.1; -  
PIR; T52378; T52378.  
InterPro: IPR005609; Sec61 beta.  
Pfam; PF03911; Sec61 beta; 1.  
Protein transport; Translocation; Transmembrane;  
Endoplasmic reticulum.  
DOMAIN 1 55 CYTOPLASMIC (POTENTIAL).  
TRANSMEM 56 76 POTENTIAL.  
SQ SEQUENCE 82 AA; 8217 MW; 5CA3EC8318EFD2F CRC64;

Query Match 52.4%; Score 22; DB 1; Length 82;  
Best Local Similarity 50.0%; Pred. No. 4.3e+02;  
Matches 4; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 1 NVLMAMNII 8  
DB 56 NVVLIMSI 63

RESULT 31  
PAK1\_SYNY3  
ID PAK1\_SYNY3 STANDARD; PRT; 86 AA.  
AC P727L2;  
DT 01-NOV-1997 (Rel. 35, Created)  
DT 01-NOV-1997 (Rel. 35, Last sequence update)  
DT 28-FEB-2003 (Rel. 41, Last annotation update)  
DE Photosystem I reaction center subunit peak I precursor (Photosystem I  
subunit X 1).  
GN PSAXI OR PSAX OR SSR0390.  
OS Synechocystis sp. (strain PCC 6803).  
OC Bacteria; Cyanobacteria; Chroococcales; Synechocystis.  
OX NCBI\_TaxID=1148;  
RN [1]  
RC SEQUENCE FROM N.A.  
RX MEDLINE=97061201; PubMed=8905231;  
RA Kaneko T., Sato S., Kotani H., Tanaka A., Asamizu E., Nakamura Y.,  
Miyajima N., Hiroseawa M., Sugliura M., Sasamoto S., Kimura T.,  
Hosouchi T., Matsuno A., Muraki A., Nakazaki N., Naruo K.,  
Okumura S., Shimpo S., Takeuchi C., Wada T., Watanabe A.,  
Okamura M., Yasuda M., Tabata S.;  
RT "Sequence analysis of the genome of the unicellular cyanobacterium  
Synechocystis sp. strain PCC6803. II. Sequence determination of the  
entire genome and assignment of potential protein-coding regions."  
RL DNA Res. 3:109-136(1996).  
CC -1- FUNCTION: THE EXACT FUNCTION OF PSAX IN PHOTOSYSTEM I (PSI) IS  
NOT YET KNOWN.  
CC -1- SUBCELLULAR LOCATION: Integral membrane protein. Cellular  
thylakoid membrane (Probable).  
CC -1- SIMILARITY: Belongs to the psaG / psak family.  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
between the Swiss Institute of Bioinformatics and the EMBL outstation -  
the European Bioinformatics Institute. There are no restrictions on its  
use by non-profit institutions as long as its content is in no way

```

Query Match      52.4%; Score 22; DB 1; Length 91;
Best Local Similarity 80.0%; Pred. No. 4.7e+02;
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY      4 MAMNI 8
      ||||:
Db      1 MAMNL 5

RESULT 33
YC23_METJA
ID_YC23_METJA      STANDARD;      PRT;      92 AA.
AC_Q58620;
DT_30-MAY-2000 (Rel. 39, Created)
DT_30-MAY-2000 (Rel. 39, Last sequence update)
DT_16-OCT-2001 (Rel. 40, Last annotation update)
DE_Hypothetical protein MJ1223.
GN_MJ1223.
OS_Methanococcus jannaschii.
OC_Archaea; Euryarchaeota; Methanococci; Methanococcales;
OC_Methanocaldococcaceae; Methanocaldococcus.
OX_NCBI_TaxID=2190;
[1]
RP_SEQUENCE FROM N.A.
RC_STRAIN=JAL-1 / DSM 2661 / ATCC 43067;
RX_MEDLINE=96337999; PubMed=8698087;
RA_Bult C.J., White O., Olsen G.J., Zhou L., Fleischmann R.D.,
RA_Sutton G.G., Blake J.A., FitzGerald L.M., Clayton R.A., Gocayne J.D.,
RA_Kerlavage A.R., Dougherty B.A., Tomb J.-F., Adams M.D., Reich C.I.,
RA_Overbeek R., Kirkness E.F., Weinstock K.G., Merrick J.M., Glodek A.,
RA_Scott J.L., Geoghegan N.S.M., Weidman J.F., Fuhrmann J.L., Nguyen D.,
RA_Utterback T.R., Kelley J.M., Peterson J.D., Sadow P.W., Hanna M.C.,
RA_Cotton M.D., Roberts K.M., Hurst M.A., Kane B.P., Borodovsky M.,
RA_Klenk H.-P., Fraser C.M., Smith H.O., Woese C.R., Venter J.C.;
RA_The complete genome sequence of the methanogenic archaeon, Methanococcus
RT_Jannaschii.;
RT_Science 273:1058-1073 (1996).
RL_SEQUENCE OF 20-28, AND CHARACTERIZATION.
CC_-1- SUBCELLULAR LOCATION: Integral membrane protein. (Potential).
CC_-1- SIMILARITY: TO M.THERMOAUTOTROPHICUM MTH1250.
CC-----
CC>This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC-----
DR_EMBL; U67563; AAB99235.1; -.
DR_PIR; F64452; F64452.
DR_TIGR; MJ1223; -.
KW_Hypothetical protein; Transmembrane; Complete proteome.
FT_TRANSMEM 1 21 POTENTIAL.
FT_TRANSMEM 30 50 POTENTIAL.
FT_TRANSMEM 62 82 POTENTIAL.
SQ_SEQUENCE 92 AA; 9739 MW; DDBCA78E28CCF164 CRC64;

Query Match      52.4%; Score 22; DB 1; Length 92;
Best Local Similarity 55.6%; Pred. No. 4.8e+02;
Matches 5; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY      1 NVLMAMNII 9
      ||||:
Db      31 NVLVGESII 39

RESULT 34
C553_HELPY
ID_C553_HELPY      STANDARD;      PRT;      96 AA.
AC_Q25825;
DT_28-FEB-2003 (Rel. 41, Created)
DT_28-FEB-2003 (Rel. 41, Last sequence update)
DT_28-FEB-2003 (Rel. 41, Last annotation update)

```

---

```

DE GN. Cytochrome C-553 precursor (C553).
OS HP1227.
OC_Helicobacter pylori (Campylobacter pylori).
OC_Bacteria; Proteobacteria; Epsilonproteobacteria; Campylobacteriales;
OC_Helicobacteraceae; Helicobacter.
OX_NCBI_TaxID=210;
[1]
RP_SEQUENCE FROM N.A.
RC_STRAIN=26695 / ATCC 700392;
RX_MEDLINE=97394467; PubMed=9252185;
RA_Tomb J.-P., White O., Kerlavage A.R., Clayton R.A., Sutton G.G.,
RA_Fleischmann R.D., Ketchum K.A., Klenk H.-P., Gill S., Dougherty B.A.,
RA_Nelson K., Quackenbush J., Zhou L., Kirkness E.F., Peterson S.,
RA_Loftus B., Richardson D., Dodson R., Khalak H.G., Glodek A.,
RA_McKenney K., FitzGerald L.M., Lee N., Adams M.D., Hickey E.K.,
RA_Berg D.E., Gocayne J.D., Utterback T.R., Peterson J.D., Kelley J.M.,
RA_Cotton M.D., Weidman J.M., Fujii C., Bowman C., Watthey L., Wallin E.,
RA_Hayes W.S., Borodovsky M., Karp P.D., Smith H.O., Fraser C.M.,
RA_Venter J.C.;
RA_The complete genome sequence of the gastric pathogen Helicobacter
RT_pylori.;
RL_Nature 388:539-547 (1997).
[2]
RP_SEQUENCE OF 20-28, AND CHARACTERIZATION.
RX_MEDLINE=20536487; PubMed=10965034;
RA_Koyanagi S., Nagata K., Tamura T., Tsukita S., Sone N.;
RA_Purification and characterization of cytochrome c-553 from
RT_Helicobacter pylori.;
RL_J Biochem. 128:371-375 (2000).
CC_-1- FUNCTION: Natural electron acceptor for a formate dehydrogenase.
CC_-1- SUBCELLULAR LOCATION: Periplasmic.
CC_-1- PFM: BINDS ONE HEME GROUP PER MOLECULE.
CC_-1- SIMILARITY: BELONGS TO THE CLASS I CYTOCHROME C FAMILY.
CC-----
CC>This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC-----
DR_EMBL; A8000628; AAD08272.1; -.
DR_PIR; C64673; C64673.
DR_HSSP; P04032; 2DVH.
DR_TIGR; HP1227; -.
DR_InterPro; IPR003088; Cyt_C1.
DR_InterPro; IPR002329; Cyt_C1C.
DR_InterPro; IPR000345; CytC_heme_bind.
DR_Pfam; PF00034; cytochrome c; 1.
DR_PRINTS; PR00605; CYTOCHROME C.
DR_ProDom; PD004020; Cyt_C_bact; 1.
DR_PROSITE; PS00190; CYTOCHROME C; 1.
KW_Electron transport; Heme; Signal; Periplasmic; Complete proteome.
FT_SIGNAL 1 19
FT_CHAIN 20 96 CYTOCHROME C-553.
FT_BINDING 29 29 HEME (COVALENT).
FT_BINDING 32 32 HEME (COVALENT).
FT_METAL 33 33 IRON (HEME AXIAL LIGAND) (BY
FT_METAL 73 73 IRON (HEME AXIAL LIGAND) (BY
FT_SIMILARITY.
SQ_SEQUENCE 96 AA; 10346 MW; C6EEFEA2AE5B8428 CRC64;

Query Match      52.4%; Score 22; DB 1; Length 96;
Best Local Similarity 37.5%; Pred. No. 5e+02;
Matches 3; Conservative 4; Mismatches 1; Indels 0; Gaps 0;

QY      2 VLMAMNII 9
      ||||:
Db      4 VINALGVL 11

```

Tue Feb 17 11:55:47 2004

```

RX MEDLINE=99098146; PubMed=9881471;
RA Ursing B.M., Arnason U.;
RT "Analyses of mitochondrial genomes strongly support a hippopotamus-
RT whale clade.";
RL Proc. R. Soc. Lond., B, Biol. Sci. 265:2251-2255(1998).
CC -1- CATALYTIC ACTIVITY: NADH + ubiquinone = NAD(+) + ubiquinol.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; AJ010957; CAA09436.1; -.
CC InterPro; IPR001133; Oxidored_4L.
CC DR InterPro; IPR001133; Oxidored_4L.
CC DR Pfam; PF00420; Oxidored_g2; 1.
CC DR ProDom; PD000359; Oxidored_4L; 1.
CC KW Oxidoreductase; NAD; Ubiquinone; Mitochondrion.
CC SEQUENCE 98 AA; 10776 MW; 5F095988CE0622F0 CRC64;
SQ
Query Match 52.4%; Score 22; DB 1; Length 98;
Best Local Similarity 71.4%; Pred. No. 5.1e+02; Indels 0; Gaps 0;
Matches 5; Conservative 1; Mismatches 1;
QY 3 LMANNNII 9
Db 3 LVYNNII 9
RESULT 37
NULM_HYLLA STANDARD; PRT; 98 AA.
ID NULM_HYLLA
AC Q95709;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 01-NOV-1997 (Rel. 35, Last annotation update)
DE NADH-ubiquinone oxidoreductase chain 4L (EC 1.6.5.3).
GN MTND4L OR ND4L OR NADH4L.
OS Hylobates lar (Common gibbon).
OG Mitochondrion.
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hylobatidae; Hylobates.
OC NCBI_TaxID=9580;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Egter;
RA Arnason U., Gullberg A., Xu X.;
RT "A complete mitochondrial DNA molecule of the white-handed gibbon,
RT Hylobates lar, and comparison among individual mitochondrial genes of
RT all hominoid genera.";
RL Hereditas 124:185-189(1996).
CC -1- CATALYTIC ACTIVITY: NADH + ubiquinone = NAD(+) + ubiquinol.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; X99256; CAA67636.1; -.
CC PIR; T11841; T11841.
CC DR InterPro; IPR001133; Oxidored_4L.
CC DR InterPro; IPR001133; Oxidored_4L.
CC DR Pfam; PF00420; Oxidored_g2; 1.
CC DR ProDom; PD000359; Oxidored_4L; 1.
CC KW Oxidoreductase; NAD; Ubiquinone; Mitochondrion.
CC SEQUENCE 98 AA; 10739 MW; 158CEBABC1614E67 CRC64;
SQ
Query Match 52.4%; Score 22; DB 1; Length 96;
Best Local Similarity 50.0%; Pred. No. 5e+02; Indels 0; Gaps 0;
Matches 4; Conservative 2; Mismatches 2;
QY 2 VLMMNNII 9
Db 57 ILMTLRII 64
RESULT 36
NULM_HIPAM STANDARD; PRT; 98 AA.
ID NULM_HIPAM
AC Q32Y3;
DT 30-MAY-2000 (Rel. 39, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE NADH-ubiquinone oxidoreductase chain 4L (EC 1.6.5.3).
GN MTND4L OR ND4L OR NADH4L.
OS Hippopotamus amphibius (Hippopotamus).
OG Mitochondrion.
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Hippopotamidae; Hippopotamus.
OC NCBI_TaxID=9633;
RN [1]
RP SEQUENCE FROM N.A.

```

```

Query Match      52.4%; Score 22; DB 1; Length 98;
Best Local Similarity 57.1%; Pred. No. 5.1e+02;
Matches 4; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 NVLMANN 7
DB 44 STLWALN 50

RESULT 38
NULM_PIG
ID NULM_PIG STANDARD; PRT; 98 AA.
AC P56632;
DT 15-DEC-1998 (Rel. 37, Created)
DT 15-DEC-1998 (Rel. 37, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE NADH-ubiquinone oxidoreductase chain 4L (EC 1.6.5.3).
GN MTND4L OR ND4L OR NADH4L.
OS Sus scrofa (Pig).
OG Mitochondrion.
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
OX NCBI_TaxID=9823;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=98404150; PubMed=9732457;
RA Ursing B.M., Arnason U.;
RT "The complete mitochondrial DNA sequence of the pig (Sus scrofa).";
RL J. Mol. Evol. 47:302-306(1998).
RN [2]
RP SEQUENCE FROM N.A.
RX STRAIN=Landrace; TISSUE=Heart;
RC MEDLINE=99365306; PubMed=10433971;
RA Lin C.S., Sun Y.L., Liu C.Y., Yang P.C., Chang L.C., Cheng I.C.,
RA Mao S.J.T., Huang M.C.;
RT "Complete nucleotide sequence of pig (Sus scrofa) mitochondrial genome
and dating evolutionary divergence within artiodactyla.";
RL Gene 236:107-114(1999).
CC -I- CATALYTIC ACTIVITY: NADH + ubiquinone = NAD(+) + ubiquinol.
-----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
-----
DR EMBL; X55514; CAA39125.1; -.
DR PIR; S14206; S14206.
DR InterPro; IPR001133; Oxidored_4L.
DR Pfam; PF00420; oxidored_q2; 1.
KW Oxidoreductase; NAD; Ubiquinone; Mitochondrion.
SQ SEQUENCE 98 AA; 10965 MW; 498EB65E8AD9AE05 CRC64;

Query Match      52.4%; Score 22; DB 1; Length 98;
Best Local Similarity 33.3%; Pred. No. 5.1e+02;
Matches 3; Conservative 5; Mismatches 1; Indels 0; Gaps 0;

QY 1 NVLMANNII 9
DB 54 NLLSNNLI 62

RESULT 40
Y233_MYCGE
ID Y233_MYCGE STANDARD; PRT; 99 AA.
AC P47475;
DT 01-FEB-1996 (Rel. 33, Created)
DT 01-FEB-1996 (Rel. 33, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Hypothetical protein MG233.
GN MG233.
OS Mycoplasma genitalium.
OC Bacteria; Firmicutes; Mollicutes; Mycoplasmataceae; Mycoplasma.
OX NCBI_TaxID=2097;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC 33530 / G-37;
RX MEDLINE=96026346; PubMed=7569993;
RA Fraser C.M., Gocayne J.D., White O., Adams M.D., Clayton R.A.,
RA Fleischmann R.D., Bult C.J., Kerlavage A.R., Sutton G., Kelley J.M.,
RA Fritchman J.L., Weidman J.F., Small K.V., Sandusky M., Fuhrmann J.L.,
RA Nguyen D.T., Uterback T.R., Saudek D.M., Phillips C.A., Merrick J.M.,
RA Tomb J.-P., Dougherty B.A., Bost K.F., Hu P.-C., Lucier T.S.,
RA Peterson S.N., Smith H.O., Hutchison C.A. III, Venter J.C.;
RT "The minimal gene complement of Mycoplasma genitalium.";
RL Science 270:397-403(1995).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC 33530 / G-37;
RX MEDLINE=94075230; PubMed=8253680;
RA Peterson S.N., Hu P.-C., Bost K.F., Hutchison C.A. III;
RT "A survey of the Mycoplasma genitalium genome by using random
sequencing.";
RL J. Bacteriol. 175:7918-7930(1993).

```

Tue Feb 17 11:55:47 2004

```
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; D13958; -; NOT_ANNOTATED_CDS.
CC DR EMBL; AE000153; AAC73568.1; -.
CC DR EMBL; U82664; AAB40220.1; -.
CC DR PIR; A64777; A64777.
CC DR EcoGene; EG12830; ybmA.
CC KW Hypothetical protein; Transmembrane; Complete proteome.
CC SQ SEQUENCE 53 AA; 6025 MW; 3200E1AC314B9D89 CRC64;
CC
CC Query Match 50.0%; Score 21; DB 1; Length 53;
CC Best Local Similarity 50.0%; Pred. NO. 4.5e+02;
CC Matches 4; Conservative 3; Mismatches 1; Indels 0; Gaps 0;
CC
CC QY 2 VLMAMNII 9
CC ||| :|:|:
CC DB 30 VIRALDIV 37
CC
CC RESULT 42
CC Y546 METJA STANDARD; PRT; 56 AA.
CC ID Y546 METJA
CC AC Q57966;
CC DT 01-NOV-1997 (Rel. 35, Created)
CC DT 01-NOV-1997 (Rel. 35, Last sequence update)
CC DT 16-OCT-2001 (Rel. 40, Last annotation update)
CC DE Hypothetical protein MJ0546.
CC GN MJ0546
CC OS Methanococcus jannaschii.
CC OC Archaea; Euryarchaeota; Methanococci; Methanococcales;
CC OC Methanocaldococcaceae; Methanocaldococcus.
CC OX NCBI_TaxID=2190;
CC RN [1]
CC RP SEQUENCE FROM N.A.
CC RC STRAIN=JAL-1 / DSM 2661 / ATCC 43067;
CC RX MEDLINE=96337999; PubMed=8688087;
CC RA Bult C.J., White O., Olsen G.J., Zhou L., Fleischmann R.D.,
CC RA Sutton G.G., Blake J.A., FitzGerald L.M., Clayton R.A., Gocayne J.D.,
CC RA Karlavage A.R., Dougherty B.A., Tomb J.-P., Adams M.D., Reich C.I.,
CC RA Overbeek R., Kirkness E.F., Weinstock K.G., Merrick J.M., Glodek A.,
CC RA Scott J.L., Geoghegan N.S.M., Weidman J.D., Sadow F.W., Nguyen D.,
CC RA Uitterback T.R., Kelley J.M., Peterson J.D., Kaine B.P., Borodovsky M.,
CC RA Cotton M.D., Roberts K.M., Hurst M.A., Kaine B.P., Borodovsky M.,
CC RA Klenk H.-P., Fraser C.M., Smith H.O., Woese C.R., Venter J.C.;
CC RA "Complete genome sequence of the methanogenic archaeon, Methanococcus
CC jannaschii.";
CC RT Science 273:1058-1073 (1996).
CC RL
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; U67504; AAB98548.1; -.
CC DR PIR; B64368; B64368.
CC DR TIGR; MJ0546; -.
CC KW Hypothetical protein; Transmembrane; Complete proteome.
CC FT TRANSMEM 2 22 POTENTIAL.
CC FT DOMAIN 27 47 GLO-RICH.
CC SQ SEQUENCE 56 AA; 6460 MW; 56CBFC395548BA5D CRC64;
CC
CC Query Match 50.0%; Score 21; DB 1; Length 56;
CC Best Local Similarity 22.2%; Pred. No. 4.7e+02;
CC
```

```
CC -----
CC -1- SIMILARITY: TO B. SUBTILIS YSXB.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; U39703; AAC71454.1; -.
CC DR EMBL; U02141; RAD12419.1; -.
CC DR PIR; G64225; G64225.
CC DR TIGR; MG233; -.
CC DR Pfam; PF04327; DUF464; 1.
CC KW Hypothetical protein; Transmembrane; Complete proteome.
CC FT TRANSMEM 26 42 POTENTIAL.
CC SQ SEQUENCE 99 AA; 10785 MW; 5B0929DD73646B28 CRC64;
CC
CC Query Match 52.4%; Score 22; DB 1; Length 99;
CC Best Local Similarity 44.4%; Pred. NO. 5.2e+02;
CC Matches 4; Conservative 3; Mismatches 2; Indels 0; Gaps 0;
CC
CC QY 1 NVLMAMNII 9
CC ||| :|:|:
CC DB 68 NVAFAFSVI 76
CC
CC RESULT 41
CC YBAM_ECOLI STANDARD; PRT; 53 AA.
CC ID YBAM_ECOLI
CC AC P45807;
CC DT 01-NOV-1995 (Rel. 32, Created)
CC DT 01-NOV-1995 (Rel. 32, Last sequence update)
CC DT 16-OCT-2001 (Rel. 40, Last annotation update)
CC DE Hypothetical protein ybam.
CC GN YBAM OR B0466.
CC OS Escherichia coli.
CC OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
CC OC Enterobacteriaceae; Escherichia.
CC OX NCBI_TaxID=562;
CC RN [1]
CC RP SEQUENCE FROM N.A.
CC RC STRAIN=K12 / MGL655;
CC RX MEDLINE=91310687; PubMed=1856227;
CC RA Blattner F.R., Plunkett G. III, Bloch C.A., Perna N.T., Burland V.,
CC RA Riley M., Collado-Vides J., Glasner J.D., Rode C.K., Mayhew G.F.,
CC RA Gregor J., Davis N.W., Kirkpatrick H.A., Goeden M.A., Rose D.J.,
CC RA Mau B., Shao Y.;
CC RA "The complete genome sequence of Escherichia coli K-12.";
CC RT Science 277:1453-1474 (1997).
CC RL [3]
CC
CC SEQUENCE FROM N.A.
CC RC STRAIN=K12 / MGL655;
CC RX MEDLINE=97426617; PubMed=9278503;
CC RA Blattner F.R., Plunkett G. III, Bloch C.A., Perna N.T., Burland V.,
CC RA Riley M., Collado-Vides J., Glasner J.D., Rode C.K., Mayhew G.F.,
CC RA Gregor J., Davis N.W., Kirkpatrick H.A., Goeden M.A., Rose D.J.,
CC RA Mau B., Shao Y.;
CC RA "The complete genome sequence of Escherichia coli K-12.";
CC RT Science 277:1453-1474 (1997).
CC RL [3]
CC
CC SEQUENCE FROM N.A.
CC RC STRAIN=K12 / MGL655;
CC RX MEDLINE=96032851; PubMed=7567469;
CC RA Borodovsky M., McIninch J., Koonin E.V., Rudd K.E., Medigue C.,
CC RA Danchin A.;
CC RA "Detection of new genes in a bacterial genome using Markov models for
CC RT three gene classes.";
CC RL Nucleic Acids Res. 23:3554-3562 (1995).
```

Matches 2; Conservative 5; Mismatches 2; Indels 0; Gaps 0;  
 QY 1 NVLMAMNII 9  
 |:::|  
 Db 14 NIILGIKVI 22

## RESULT 43

YC33 CYAPA STANDARD; PRT; 65 AA.  
 AC P48273;  
 DT 01-FEB-1996 (Rel. 33, Created)  
 DT 01-FEB-1996 (Rel. 33, Last sequence update)  
 DT 16-OCT-2001 (Rel. 40, Last annotation update)  
 DE Hypothetical 7.6 kDa protein ycf33.  
 GN YCF33.  
 OS Cyanophora paradoxa.  
 OG Cyanelle.  
 OC Eukaryota; Glaucocystophyceae; Cyanophoraceae; Cyanophora.  
 OX NCBI\_TaxID=2762;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=UTEX LB 555 / Pringsheim;  
 RA Stirewalt V.L., Michalowski C.B., Loeffelhardt W., Bohnert H.J.,  
 RA Bryant D.A.;  
 RT "Nucleotide sequence of the cyanelle DNA from Cyanophora paradoxa.";  
 RL Plant Mol. Biol. Rep. 13:327-332(1995).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=UTEX LB 555 / Pringsheim;  
 RA Loeffelhardt W., Stirewalt V.L., Michalowski C.B., Annarella M.,  
 RA Farley J.Y., Schluchter W.M., Chung S., Newman-Spallart C.,  
 RA Steiner J.M., Jakowitsch J., Bohnert H.J., Bryant D.A.;  
 RT "The complete sequence of the cyanelle genome of Cyanophora paradoxa:  
 the genetic complexity of a primitive plastid.";  
 RL (In) Schenk H.E.A., Herrmann R., Jeon K.W., Mueller N.E.,  
 RL Schweinmler W. (eds.);  
 RL Eukaryotism and Symbiosis, pp.40-48, Springer-Verlag, Heidelberg  
 (1997).  
 CC -1- SIMILARITY: BELONGS TO THE YCF33 FAMILY.

-----  
 This SWISS-PROT entry is copyright. It is produced through a collaboration  
 between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 the European Bioinformatics Institute. There are no restrictions on its  
 use by non-profit institutions as long as its content is in no way  
 modified and this statement is not removed. Usage by and for commercial  
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 -----

EMBL; U30821; AAA81285.1; -;  
 PIR; T06942; T06942.  
 KW Cyanelle; Hypothetical protein.  
 SQ SEQUENCE 65 AA; 7635 MW; EAFDDC8B6233CE3 CRC64;

Query Match 50.0%; Score 21; DB 1; Length 65;  
 Best Local Similarity 57.1%; Pred. No. 5.6e+02;  
 Matches 4; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 2 VLMAMNI 8  
 |:::|  
 Db 50 VLMFLNL 56

## RESULT 44

CSPA\_RICCN STANDARD; PRT; 70 AA.  
 AC Q92GV1;  
 DT 28-FEB-2003 (Rel. 41, Created)  
 DT 28-FEB-2003 (Rel. 41, Last sequence update)  
 DT 28-FEB-2003 (Rel. 41, Last annotation update)  
 DE Cold shock-like protein cspa.  
 GN CSPA OR RC1021.  
 OS Rickettsia conorii.

OC Bacteria; Proteobacteria; Alphaproteobacteria; Rickettsiales;  
 OC Rickettsiaceae; Rickettsiae; Rickettsia.  
 OX NCBI\_TaxID=781;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=Mailish 7;  
 RX MEDLINE=21442074; PubMed=11557893;  
 RA Ogata H., Audic S., Renesto-Audiffren P., Fournier P.-E., Barbe V.,  
 RA Samson D., Roux V., Cossart P., Weissenbach J., Claverie J.-M.,  
 RA Raoult D.;  
 RT "Mechanisms of evolution in Rickettsia conorii and R. prowazekii.";  
 RL Science 293:2093-2098(2001).  
 CC -1- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).  
 CC -1- SIMILARITY: BELONGS TO THE COLD-SHOCK DOMAIN (CSD) FAMILY.  
 -----  
 This SWISS-PROT entry is copyright. It is produced through a collaboration  
 between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 the European Bioinformatics Institute. There are no restrictions on its  
 use by non-profit institutions as long as its content is in no way  
 modified and this statement is not removed. Usage by and for commercial  
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 -----

EMBL; AK008653; AAL03559.1; -;  
 PIR; E97827; E97827.  
 DR InterPro; IPR002059; Cold\_shock.  
 DR Pfam; PF00313; CSD; 1.  
 DR PRINTS; PR00050; COLDSHOCK.  
 DR ProDom; PD000621; Cold\_shock; 1.  
 DR SMART; SM00357; CSP; 1.  
 DR PROSITE; PS00352; COLD\_SHOCK; 1.  
 KW Transcription regulation; DNA-binding; Activator; Complete proteome.  
 FT DOMAIN 7 67 CSD.  
 SQ SEQUENCE 70 AA; 7771 MW; 822E9714229FE03A CRC64;

Query Match 50.0%; Score 21; DB 1; Length 70;  
 Best Local Similarity 66.7%; Pred. No. 6e+02;  
 Matches 4; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 4 MAMNII 9  
 |:::|  
 Db 1 MATNIV 6

## RESULT 45

CSPA\_RICPR STANDARD; PRT; 70 AA.  
 AC Q9ZCF9;  
 DT 30-MAY-2000 (Rel. 39, Created)  
 DT 30-MAY-2000 (Rel. 39, Last sequence update)  
 DT 28-FEB-2003 (Rel. 41, Last annotation update)  
 DE Cold shock-like protein cspa.  
 GN CSPA OR RP670.  
 OS Rickettsia prowazekii.  
 OC Bacteria; Proteobacteria; Alphaproteobacteria; Rickettsiales;  
 OC Rickettsiaceae; Rickettsiae; Rickettsia.  
 OX NCBI\_TaxID=782;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=Madrid E;  
 RX MEDLINE=99039499; PubMed=9823893;  
 RA Andersson S.G.E., Zomorodipour A., Andersson J.O.,  
 RA Scharitz-Ponten T., Alsmark U.C.M., Podowski R.M., Naeslund A.K.,  
 RA Eriksson A.-S., Winkler H.H., Kurland C.G.;  
 RT "The genome sequence of Rickettsia prowazekii and the origin of  
 mitochondria.";  
 RL Nature 396:133-140(1998).  
 CC -1- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).  
 CC -1- SIMILARITY: BELONGS TO THE COLD-SHOCK DOMAIN (CSD) FAMILY.

-----  
 This SWISS-PROT entry is copyright. It is produced through a collaboration  
 between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 the European Bioinformatics Institute. There are no restrictions on its

Tue Feb 17 11:55:47 2004

CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----

DR ENBL; AJ235272; CAA15108.1; --  
DR PIR; B71673; B71673.  
DR HSPP; P41016; IC90.  
DR InterPro; IPR002059; Cold\_shock.  
DR Pfam; PF00313; CSD; 1.  
DR PRINTS; PR00050; COLDSHOCK.  
DR ProDom; PD000621; Cold\_shock; 1.  
DR SMART; SM00357; CSP; 1.  
DR PROSITE; PS00352; COLD\_SHOCK; 1.  
KW Transcription regulation; DNA-binding; Activator; Complete proteome.  
FT DOMAIN 7  
SQ SEQUENCE 70 AA; 7785 MW; F49FBF2FB344903E CRC64;

Query Match 50.0%; Score 21; DB 1; Length 70;  
Best Local Similarity 66.7%; Pred. No. 6e+02;  
Matches 4; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

OY 4 MAMNII 9  
Db 1 MATNIV 6

Search completed: February 17, 2004, 10:57:01  
Job time : 4.85149 secs



GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: February 17, 2004, 10:50:13 ; Search time 6.50495 Seconds  
(without alignments)  
58.540 Million cell updates/sec

Title: US-09-900-147-2  
Perfect score: 42  
Sequence: 1 NVLMANNII 9

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 328717 seqs, 42310858 residues

Total number of hits satisfying chosen parameters: 228043

Minimum DB seq length: 0  
Maximum DB seq length: 100

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Issued Patents AA:\*  
1: /cgn2\_6/ptodata/1/iaa/5A COMB.pep.\*  
2: /cgn2\_6/ptodata/1/iaa/5B COMB.pep.\*  
3: /cgn2\_6/ptodata/1/iaa/6A COMB.pep.\*  
4: /cgn2\_6/ptodata/1/iaa/6B COMB.pep.\*  
5: /cgn2\_6/ptodata/1/iaa/PCTUS COMB.pep.\*  
6: /cgn2\_6/ptodata/1/iaa/backfiles1.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	42	100.0	9	3	US-09-308-935-2
2	42	100.0	16	3	US-09-308-935-5
3	42	100.0	17	2	US-08-428-131-13
4	42	100.0	17	3	US-09-078-596-13
5	42	100.0	19	3	US-09-308-935-3
6	42	100.0	19	3	US-09-308-935-16
7	42	100.0	20	3	US-09-308-935-4
8	42	100.0	30	3	US-09-308-935-6
9	42	100.0	37	3	US-09-308-935-1
10	42	100.0	72	2	US-08-428-131-11
11	42	100.0	72	3	US-09-078-596-11
12	42	100.0	74	4	US-08-894-139-10
13	38	90.5	19	3	US-09-308-935-15
14	36	85.7	19	3	US-09-308-935-17
15	34	81.0	14	3	US-09-308-935-11
16	33	78.6	63	1	US-08-194-338-14
17	29	69.0	88	4	US-08-936-165A-513
18	29	69.0	89	4	US-09-134-001C-3697
19	29	69.0	100	4	US-09-134-001C-3434
20	27	64.3	60	4	US-09-107-532A-3973
21	27	64.3	92	4	US-09-107-532A-4069
22	25	59.5	31	2	US-08-031-538-68
23	25	59.5	62	4	US-09-134-001C-3106
24	25	59.5	66	4	US-09-205-258-1030
25	25	59.5	84	4	US-09-205-258-1032
26	25	59.5	93	4	US-09-134-001C-4976
27	25	59.5	95	4	US-09-328-352-8011

28	57.1	37	1	US-07-626-618A-6	Sequence 6, Appli
29	57.1	37	1	US-07-928-611-6	Sequence 6, Appli
30	57.1	37	1	US-08-333-977-6	Sequence 6, Appli
31	57.1	37	2	US-08-487-811A-6	Sequence 6, Appli
32	57.1	37	3	US-09-060-694-6	Sequence 6, Appli
33	57.1	37	4	US-09-378-074-6	Sequence 6, Appli
34	57.1	37	5	PCT-US93-07370-6	Sequence 6, Appli
35	57.1	40	3	US-09-215-221-18	Sequence 18, Appli
36	57.1	69	4	US-09-252-991A-30061	Sequence 30061, A
37	57.1	76	4	US-09-134-001C-4499	Sequence 4499, Ap
38	57.1	77	4	US-09-328-352-5816	Sequence 5816, Ap
39	57.1	85	3	US-09-103-478-27	Sequence 27, Appli
40	57.1	85	4	US-09-193-931C-27	Sequence 27, Appli
41	57.1	87	1	US-08-289-247B-4	Sequence 4, Appli
42	57.1	87	1	US-08-725-531-4	Sequence 4, Appli
43	57.1	87	2	US-08-738-127-4	Sequence 4, Appli
44	57.1	87	2	US-09-213-392-4	Sequence 4, Appli
45	57.1	87	2	US-09-083-661-4	Sequence 4, Appli

## ALIGNMENTS

RESULT 1  
US-09-308-935-2  
; Sequence 2, Application US/09308935  
; Patent No. 6268334  
; GENERAL INFORMATION:  
; APPLICANT: La Thangue, Nicholas B  
; APPLICANT: Bandara, Lasantha R  
; TITLE OF INVENTION: Peptide antagonists of DP transcription factors  
; FILE REFERENCE: 620-67  
; CURRENT APPLICATION NUMBER: US/09/308,935  
; CURRENT FILING DATE: 1999-05-27  
; EARLIER APPLICATION NUMBER: PCT/GB97/03506  
; EARLIER FILING DATE: 1997-12-22  
; EARLIER APPLICATION NUMBER: GB 9626589.7  
; EARLIER FILING DATE: 1996-12-20  
; NUMBER OF SEQ ID NOS: 18  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 2  
; LENGTH: 9  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Description of Artificial Sequence: Synthetic  
US-09-308-935-2

Query Match 100.0%; Score 42; DB 3; Length 9;  
Best Local Similarity 100.0%; Pred. No. 2.5e+05;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 NVLMANNII 9  
DB 1 NVLMANNII 9

RESULT 2  
US-09-308-935-5  
; Sequence 5, Application US/09308935  
; Patent No. 6268334  
; GENERAL INFORMATION:  
; APPLICANT: La Thangue, Nicholas B  
; APPLICANT: Bandara, Lasantha R  
; TITLE OF INVENTION: Peptide antagonists of DP transcription factors  
; FILE REFERENCE: 620-67  
; CURRENT APPLICATION NUMBER: US/09/308,935  
; CURRENT FILING DATE: 1999-05-27  
; EARLIER APPLICATION NUMBER: PCT/GB97/03506  
; EARLIER FILING DATE: 1997-12-22  
; EARLIER APPLICATION NUMBER: GB 9626589.7  
; EARLIER FILING DATE: 1996-12-20  
; NUMBER OF SEQ ID NOS: 18

Tue Feb 17 11:55:47 2004

us-09-900-147-2.rai

```

; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 5
; LENGTH: 16
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Synthetic
US-09-308-935-5

Query Match      100.0%; Score 42; DB 3; Length 16;
Best Local Similarity 100.0%; Pred. No. 0.013; 0; Indels 0; Gaps 0;
Matches 9; Conservative 0; Mismatches 0;

Qy 1 NVLMAMNII 9
Db 7 NVLMAMNII 15

RESULT 3
US-08-428-131-13
; Sequence 13, Application US/08428131
; Patent No. 5863757
; GENERAL INFORMATION:
; APPLICANT: La Thangue, Nicholas Barrie
; TITLE OF INVENTION: Transcription Factor DP-1
; NUMBER OF SEQUENCES: 14
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Nixon & Vanderhye
; STREET: 1100 No. 5863757th Glebe Road, 8th Floor
; CITY: Arlington
; STATE: Virginia
; COUNTRY: U.S.A.
; ZIP: 22201-4714
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25 (EPO)
; CURRENT APPLICATION NUMBER: US/08/428.131
; FILING DATE: 23-JUN-1995
; CLASSIFICATION: 514
; ATTORNEY/AGENT INFORMATION:
; NAME: Arthur R. Crawford
; REGISTRATION NUMBER: 25,327
; REFERENCE/DOCKET NUMBER: 117-181
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703) 816-4000
; TELEFAX: (703) 816-4100
; INFORMATION FOR SEQ ID NO: 13:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 17 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-09-078-596-13

Query Match      100.0%; Score 42; DB 3; Length 17;
Best Local Similarity 100.0%; Pred. No. 0.013; 0; Indels 0; Gaps 0;
Matches 9; Conservative 0; Mismatches 0;

Qy 1 NVLMAMNII 9
Db 8 NVLMAMNII 16

RESULT 5
US-09-308-935-3
; Sequence 3, Application US/09308935
; Patent No. 6268334
; GENERAL INFORMATION:
; APPLICANT: La Thangue, Nicholas B
; APPLICANT: Bandara, Laseantha R
; TITLE OF INVENTION: Peptide antagonists of DP transcription factors
; FILE REFERENCE: 620-67
; CURRENT APPLICATION NUMBER: US/09/308,935
; CURRENT FILING DATE: 1999-05-27
; EARLIER APPLICATION NUMBER: PCT/GB97/03506
; EARLIER FILING DATE: 1997-12-22
; EARLIER APPLICATION NUMBER: GB 9626589.7
; EARLIER FILING DATE: 1996-12-20
; NUMBER OF SEQ ID NOS: 18
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 3
; LENGTH: 19
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Synthetic
US-09-308-935-3

Query Match      100.0%; Score 42; DB 3; Length 19;
Best Local Similarity 100.0%; Pred. No. 0.015; 0; Indels 0; Gaps 0;
Matches 9; Conservative 0; Mismatches 0;

Qy 1 NVLMAMNII 9
Db 8 NVLMAMNII 16

RESULT 4
US-09-078-596-13
; Sequence 13, Application US/09078596
; Patent No. 6150116
; GENERAL INFORMATION:
; APPLICANT: La Thangue, Nicholas Barrie

```

Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 NVLMANNII 9  
|||||||  
Db 9 NVLMANNII 17

## RESULT 6

US-09-308-935-16  
; Sequence 16, Application US/09308935  
; Patent No. 6268334  
; GENERAL INFORMATION:  
; APPLICANT: La Thangue, Nicholas B  
; APPLICANT: Bandara, Lasantha R  
; TITLE OF INVENTION: Peptide antagonists of DP transcription factors  
; FILE REFERENCE: 620-67  
; CURRENT APPLICATION NUMBER: US/09/308,935  
; CURRENT FILING DATE: 1999-05-27  
; EARLIER APPLICATION NUMBER: PCT/GB97/03506  
; EARLIER FILING DATE: 1997-12-22  
; EARLIER APPLICATION NUMBER: GB 9626589.7  
; EARLIER FILING DATE: 1996-12-20  
; NUMBER OF SEQ ID NOS: 18  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 16  
; LENGTH: 19  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Description of Artificial Sequence: Mutant peptide  
US-09-308-935-16

Query Match 100.0%; Score 42; DB 3; Length 19;  
Best Local Similarity 100.0%; Pred. No. 0.015;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 NVLMANNII 9  
|||||||  
Db 9 NVLMANNII 17

## RESULT 7

US-09-308-935-4  
; Sequence 4, Application US/09308935  
; Patent No. 6268334  
; GENERAL INFORMATION:  
; APPLICANT: La Thangue, Nicholas B  
; APPLICANT: Bandara, Lasantha R  
; TITLE OF INVENTION: Peptide antagonists of DP transcription factors  
; FILE REFERENCE: 620-67  
; CURRENT APPLICATION NUMBER: US/09/308,935  
; CURRENT FILING DATE: 1999-05-27  
; EARLIER APPLICATION NUMBER: PCT/GB97/03506  
; EARLIER FILING DATE: 1997-12-22  
; EARLIER APPLICATION NUMBER: GB 9626589.7  
; EARLIER FILING DATE: 1996-12-20  
; NUMBER OF SEQ ID NOS: 18  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 4  
; LENGTH: 20  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Description of Artificial Sequence: Synthetic  
US-09-308-935-4

Query Match 100.0%; Score 42; DB 3; Length 20;  
Best Local Similarity 100.0%; Pred. No. 0.016;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 NVLMANNII 9  
|||||||  
Db 1 NVLMANNII 9

## RESULT 8

US-09-308-935-6  
; Sequence 6, Application US/09308935  
; Patent No. 6268334  
; GENERAL INFORMATION:  
; APPLICANT: La Thangue, Nicholas B  
; APPLICANT: Bandara, Lasantha R  
; TITLE OF INVENTION: Peptide antagonists of DP transcription factors  
; FILE REFERENCE: 620-67  
; CURRENT APPLICATION NUMBER: US/09/308,935  
; CURRENT FILING DATE: 1999-05-27  
; EARLIER APPLICATION NUMBER: PCT/GB97/03506  
; EARLIER FILING DATE: 1997-12-22  
; EARLIER APPLICATION NUMBER: GB 9626589.7  
; EARLIER FILING DATE: 1996-12-20  
; NUMBER OF SEQ ID NOS: 18  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 6  
; LENGTH: 30  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Description of Artificial Sequence: Synthetic  
US-09-308-935-6

Query Match 100.0%; Score 42; DB 3; Length 30;  
Best Local Similarity 100.0%; Pred. No. 0.025;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 NVLMANNII 9  
|||||||  
Db 5 NVLMANNII 13

## RESULT 9

US-09-308-935-1  
; Sequence 1, Application US/09308935  
; Patent No. 6268334  
; GENERAL INFORMATION:  
; APPLICANT: La Thangue, Nicholas B  
; APPLICANT: Bandara, Lasantha R  
; TITLE OF INVENTION: Peptide antagonists of DP transcription factors  
; FILE REFERENCE: 620-67  
; CURRENT APPLICATION NUMBER: US/09/308,935  
; CURRENT FILING DATE: 1999-05-27  
; EARLIER APPLICATION NUMBER: PCT/GB97/03506  
; EARLIER FILING DATE: 1997-12-22  
; EARLIER APPLICATION NUMBER: GB 9626589.7  
; EARLIER FILING DATE: 1996-12-20  
; NUMBER OF SEQ ID NOS: 18  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 1  
; LENGTH: 37  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Description of Artificial Sequence: Synthetic  
US-09-308-935-1

Query Match 100.0%; Score 42; DB 3; Length 37;  
Best Local Similarity 100.0%; Pred. No. 0.031;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 NVLMANNII 9  
|||||||  
Db 12 NVLMANNII 20

## RESULT 10

US-08-428-131-11  
; Sequence 11, Application US/08428131

Tue Feb 17 11:55:47 2004

```

; Patent No. 5861757
; GENERAL INFORMATION:
; APPLICANT: La Thangue, Nicholas Barrie
; TITLE OF INVENTION: Transcription Factor DP-1
; NUMBER OF SEQUENCES: 14
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Nixon & Vanderhye
; STREET: 1100 No. 5863757th Glebe Road, 8th Floor
; CITY: Arlington
; STATE: Virginia
; COUNTRY: U.S.A.
; ZIP: 22201-4714
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25 (EPO)
; CURRENT APPLICATION NUMBER: US/08/428,131
; FILING DATE: 23-JUN-1995
; CLASSIFICATION: 514
; ATTORNEY/AGENT INFORMATION:
; NAME: Arthur R. Crawford
; REGISTRATION NUMBER: 25,327
; REFERENCE/DOCKET NUMBER: 117-181
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703) 816-4000
; TELEFAX: (703) 816-4100
; INFORMATION FOR SEQ ID NO: 11:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 72 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-09-078-596-11

Query Match 100.0%; Score 42; DB 3; Length 72;
Best Local Similarity 100.0%; Pred. No. 0.066;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 NVLMAMNII 9
DB 15 NVLMAMNII 23

RESULT 12
US-08-894-139-10
; Sequence 10, Application US/08894139
; Patent No. 6448376
; GENERAL INFORMATION:
; APPLICANT: LA THANGUE, NICHOLAS B.
; APPLICANT: BERNARDS, RENE
; APPLICANT: HIJMAN, ELEANORE M.
; TITLE OF INVENTION: TRANSCRIPTION FACTOR E2F-5
; NUMBER OF SEQUENCES: 25
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: NIXON & VANDERHYE P.C.
; STREET: 1100 NORTH GLEBE ROAD
; CITY: ARLINGTON
; STATE: VIRGINIA
; COUNTRY: U.S.A.
; ZIP: 22201-4714
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/894,139
; FILING DATE: 13-AUG-1997
; CLASSIFICATION: 536
; ATTORNEY/AGENT INFORMATION:
; NAME: WILSON, MARY J.
; REGISTRATION NUMBER: 32,955
; REFERENCE/DOCKET NUMBER: 620-22
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703) 816-4000
; TELEFAX: (703) 816-4100
; INFORMATION FOR SEQ ID NO: 10:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 74 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; US-08-894-139-10

Query Match 100.0%; Score 42; DB 4; Length 74;
Best Local Similarity 100.0%; Pred. No. 0.068;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 NVLMAMNII 9
DB 15 NVLMAMNII 23

RESULT 11
US-09-078-596-11
; Sequence 11, Application US/09078596
; Patent No. 615016
; GENERAL INFORMATION:
; APPLICANT: La Thangue, Nicholas Barrie
; TITLE OF INVENTION: Transcription Factor DP-1
; NUMBER OF SEQUENCES: 14
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Nixon & Vanderhye
; STREET: 1100 No. 6150116th Glebe Road, 8th Floor
; CITY: Arlington
; STATE: Virginia
; COUNTRY: U.S.A.
; ZIP: 22201-4714
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25 (EPO)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/078,596
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/428,131
; FILING DATE: 23-JUN-1995

```

Db 54 NVLMAMNII 62

RESULT 13  
US-09-308-935-15  
; Sequence 15, Application US/09308935  
; Patent No. 6268334  
; GENERAL INFORMATION:  
; APPLICANT: La Thangue, Nicholas B  
; APPLICANT: Bandara, Lasantha R  
; TITLE OF INVENTION: Peptide antagonists of DP transcription factors  
; FILE REFERENCE: 620-67  
; CURRENT APPLICATION NUMBER: US/09/308,935  
; CURRENT FILING DATE: 1999-05-27  
; EARLIER APPLICATION NUMBER: PCT/GB97/03506  
; EARLIER FILING DATE: 1997-12-22  
; EARLIER APPLICATION NUMBER: GB 9626589.7  
; EARLIER FILING DATE: 1996-12-20  
; NUMBER OF SEQ ID NOS: 18  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 15  
; LENGTH: 19  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Description of Artificial Sequence: Mutant peptide  
US-09-308-935-15

Query Match 90.5%; Score 38; DB 3; Length 19;  
Best Local Similarity 88.9%; Pred. No. 0.097; 1; Indels 0; Gaps 0;  
Matches 8; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 NVLMAMNII 9  
| | | | |  
Db 9 NALMAMNII 17

RESULT 14  
US-09-308-935-17  
; Sequence 17, Application US/09308935  
; Patent No. 6268334  
; GENERAL INFORMATION:  
; APPLICANT: La Thangue, Nicholas B  
; APPLICANT: Bandara, Lasantha R  
; TITLE OF INVENTION: Peptide antagonists of DP transcription factors  
; FILE REFERENCE: 620-67  
; CURRENT APPLICATION NUMBER: US/09/308,935  
; CURRENT FILING DATE: 1999-05-27  
; EARLIER APPLICATION NUMBER: PCT/GB97/03506  
; EARLIER FILING DATE: 1997-12-22  
; EARLIER APPLICATION NUMBER: GB 9626589.7  
; EARLIER FILING DATE: 1996-12-20  
; NUMBER OF SEQ ID NOS: 18  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 17  
; LENGTH: 19  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Description of Artificial Sequence: Mutant peptide  
US-09-308-935-17

Query Match 85.7%; Score 36; DB 3; Length 19;  
Best Local Similarity 88.9%; Pred. No. 0.25; 1; Indels 0; Gaps 0;  
Matches 8; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 NVLMAMNII 9  
| | | | |  
Db 9 NVRMAMNII 17

RESULT 15  
US-09-308-935-11

; Sequence 11, Application US/09308935  
; Patent No. 6268334  
; GENERAL INFORMATION:  
; APPLICANT: La Thangue, Nicholas B  
; APPLICANT: Bandara, Lasantha R  
; TITLE OF INVENTION: Peptide antagonists of DP transcription factors  
; FILE REFERENCE: 620-67  
; CURRENT APPLICATION NUMBER: US/09/308,935  
; CURRENT FILING DATE: 1999-05-27  
; EARLIER APPLICATION NUMBER: PCT/GB97/03506  
; EARLIER FILING DATE: 1997-12-22  
; EARLIER APPLICATION NUMBER: GB 9626589.7  
; EARLIER FILING DATE: 1996-12-20  
; NUMBER OF SEQ ID NOS: 18  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 11  
; LENGTH: 14  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Description of Artificial Sequence: Synthetic  
US-09-308-935-11

Query Match 81.0%; Score 34; DB 3; Length 14;  
Best Local Similarity 100.0%; Pred. No. 0.45; 0; Indels 0; Gaps 0;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 NVLMAMN 7  
| | | | |  
Db 8 NVLMAMN 14

RESULT 16  
US-08-194-338-14  
; Sequence 14, Application US/08194338  
; Patent No. 5474898  
; GENERAL INFORMATION:  
; APPLICANT: Venter, John C.  
; APPLICANT: Fraser, Claire M.  
; APPLICANT: McCombie, William R.  
; TITLE OF INVENTION: OCTOPAMINE RECEPTOR  
; NUMBER OF SEQUENCES: 16  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Knobbe, Martens, Olson and Bear  
; STREET: 620 Newport Center Drive, Sixteenth Floor  
; CITY: Newport Beach  
; STATE: CA  
; COUNTRY: USA  
; ZIP: 92660  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/194,338  
; FILING DATE: 08-FEB-1994  
; CLASSIFICATION: 435  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 07/676,174  
; FILING DATE: 28-MAR-1991  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Israelsen, Ned A.  
; REGISTRATION NUMBER: 29,655  
; REFERENCE/DOCKET NUMBER: NIH101.001DV1  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (619) 235-8550  
; TELEFAX: (619) 235-0176  
; INFORMATION FOR SEQ ID NO: 14:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 63 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: single

TOPOLOGY: linear  
MOLECULE TYPE: peptide  
HYPOTHETICAL: NO  
ANTI-SENSE: NO  
FRAGMENT TYPE: internal  
US-08-194-338-14

Query Match 78.6%; Score 33; DB 1; Length 63;  
Best Local Similarity 66.7%; Pred. No. 3; 9;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 NVLMAMNII 9  
Db 3 NVLVALNIL 11

RESULT 17  
US-08-936-165A-513

; Sequence 513, Application US/08936165A  
; Patent No. 6348582

; GENERAL INFORMATION:

; APPLICANT: Black, Michael  
; APPLICANT: Burnham, Martin  
; APPLICANT: Hodgson, John  
; APPLICANT: Knowles, David  
; APPLICANT: Lonetto, Michael  
; APPLICANT: Nicholas, Richard  
; APPLICANT: Pratt, Julie  
; APPLICANT: Reichard, Richard  
; APPLICANT: Rosenberg, Martin  
; APPLICANT: Rosenburg, Martin  
; APPLICANT: Ward, Judith

; TITLE OF INVENTION: No. 6348582el Prokaryotic Polynucleotides,  
; TITLE OF INVENTION: Polypeptides and Their Uses

; NUMBER OF SEQUENCES: 534

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: SmithKline Beecham Corporation  
; STREET: 709 Swedeland Road  
; CITY: King of Prussia  
; STATE: PA  
; COUNTRY: USA

; ZIP: 19406-0939

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Diskette

; OPERATING SYSTEM: DOS

; SOFTWARE: FASTSEQ for Windows Version 2.0

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/936,165A

; FILING DATE: 24-SEP-1997

; CLASSIFICATION: 536

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: 60/027,032

; FILING DATE: 24-SEP-1996

; ATTORNEY/AGENT INFORMATION:

; NAME: Gimmi, Edward R

; REGISTRATION NUMBER: 38,891

; REFERENCE/DOCKET NUMBER: P50549

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: 610-270-4478

; TELEFAX: 610-270-5090

; TELEX:

; INFORMATION FOR SEQ ID NO: 513:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 88 amino acids

; TYPE: amino acid

; STRANDEDNESS: single

; TOPOLOGY: linear

; MOLECULE TYPE: Protein

US-08-936-165A-513

Query Match 69.0%; Score 29; DB 4; Length 88;

Best Local Similarity 66.7%; Pred. No. 38;

Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Qy 1 NVLMAMNII 9  
Db 55 NVLVPINII 63

RESULT 18

US-09-134-001C-3697

; Sequence 3697, Application US/09134001C

; Patent No. 6380370

; GENERAL INFORMATION:

; APPLICANT: Lynn Doucette-Stamm et al

; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO STAPHYLOCOCCUS

; TITLE OF INVENTION: EPIDERMIDIS FOR DIAGNOSTICS AND THERAPEUTICS

; FILE REFERENCE: GTC-007

; CURRENT APPLICATION NUMBER: US/09/134,001C

; CURRENT FILING DATE: 1998-08-13

; PRIOR APPLICATION NUMBER: US 60/064,964

; PRIOR FILING DATE: 1997-11-08

; PRIOR APPLICATION NUMBER: US 60/055,779

; PRIOR FILING DATE: 1997-08-14

; NUMBER OF SEQ ID NOS: 5674

; SEQ ID NO 3697

; LENGTH: 89

; TYPE: PRT

; ORGANISM: Staphylococcus epidermidis

US-09-134-001C-3697

Query Match 69.0%; Score 29; DB 4; Length 89;

Best Local Similarity 85.7%; Pred. No. 38;

Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 2 VLMAMNI 8  
Db 59 VLMALNI 65

RESULT 19

US-09-134-001C-3434

; Sequence 3434, Application US/09134001C

; Patent No. 6380370

; GENERAL INFORMATION:

; APPLICANT: Lynn Doucette-Stamm et al

; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO STAPHYLOCOCCUS

; TITLE OF INVENTION: EPIDERMIDIS FOR DIAGNOSTICS AND THERAPEUTICS

; FILE REFERENCE: GTC-007

; CURRENT APPLICATION NUMBER: US/09/134,001C

; CURRENT FILING DATE: 1998-08-13

; PRIOR APPLICATION NUMBER: US 60/064,964

; PRIOR FILING DATE: 1997-11-08

; PRIOR APPLICATION NUMBER: US 60/055,779

; PRIOR FILING DATE: 1997-08-14

; NUMBER OF SEQ ID NOS: 5674

; SEQ ID NO 3434

; LENGTH: 100

; TYPE: PRT

; ORGANISM: Staphylococcus epidermidis

US-09-134-001C-3434

Query Match 69.0%; Score 29; DB 4; Length 100;

Best Local Similarity 85.7%; Pred. No. 43;

Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 2 VLMAMNI 8  
Db 49 VLMALNI 55

RESULT 20

US-09-107-532A-3973

; Sequence 3973, Application US/09107532A

; Patent No. 6583275

; GENERAL INFORMATION:

```

; APPLICANT: Lynn A Doucette-Stamm and David Bush
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO
; ENTEROCOCCUS FAECIUM FOR DIAGNOSTICS AND THERAPEUTICS
;
; NUMBER OF SEQUENCES: 7310
; CORRESPONDENCE ADDRESSES:
; ADDRESSEE: GENOME THERAPEUTICS CORPORATION
; STREET: 100 Beaver Street
; CITY: Waltham
; STATE: Massachusetts
; COUNTRY: USA
; ZIP: 02354
;
; COMPUTER READABLE FORM:
; MEDIUM TYPE: CD/ROM ISO9660
; COMPUTER: PC
; OPERATING SYSTEM: <Unknown>
; SOFTWARE: ASCII
;
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/107,532A
; FILING DATE: 30-Jun-1998
; APPLICATION NUMBER: 60/085,598
; FILING DATE: 14 May 1998
; APPLICATION NUMBER: 60/051571
; FILING DATE: July 2, 1997
;
; ATTORNEY/AGENT INFORMATION:
; NAME: Ariniello, Pamela Deneke
; REGISTRATION NUMBER: 40,489
; REFERENCE/DOCKET NUMBER: GTC-012
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (781)893-5007
; TELEFAX: (781)893-8277
;
; INFORMATION FOR SEQ ID NO: 3973:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 60 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; HYPOTHETICAL: YES
; ORIGINAL SOURCE:
; ORGANISM: Enterococcus faecium
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (B) LOCATION 1...60
; SEQUENCE DESCRIPTION: SEQ ID NO: 3973:
;
; US-09-107-532A-3973
;
; Query Match 64.3%; Score 27; DB 4; Length 60;
; Best Local Similarity 44.4%; Pred. No. 63;
; Matches 4; Conservative 4; Mismatches 1; Indels 0; Gaps 0;
;
; QY 1 NVLMAMNII 9
; |.:|.:|.:|
; Db 19 NILYSUNIL 27
;
; RESULT 21
; US-09-107-532A-4069
; Sequence 4069, Application US/09107532A
; Patent No. 6583275
; GENERAL INFORMATION:
; APPLICANT: Lynn A Doucette-Stamm and David Bush
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO
; ENTEROCOCCUS FAECIUM FOR DIAGNOSTICS AND THERAPEUTICS
;
; NUMBER OF SEQUENCES: 7310
; CORRESPONDENCE ADDRESSES:
; ADDRESSEE: GENOME THERAPEUTICS CORPORATION
; STREET: 100 Beaver Street
; CITY: Waltham
; STATE: Massachusetts
; COUNTRY: USA
; ZIP: 02354
;
; COMPUTER READABLE FORM:
; MEDIUM TYPE: CD/ROM ISO9660

```

```

; COMPUTER: PC
; OPERATING SYSTEM: <Unknown>
; SOFTWARE: ASCII
;
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/107,532A
; FILING DATE: 30-Jun-1998
; APPLICATION NUMBER: 60/085,598
; FILING DATE: 14 May 1998
; APPLICATION NUMBER: 60/051571
; FILING DATE: July 2, 1997
;
; ATTORNEY/AGENT INFORMATION:
; NAME: Ariniello, Pamela Deneke
; REGISTRATION NUMBER: 40,489
; REFERENCE/DOCKET NUMBER: GTC-012
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (781)893-5007
; TELEFAX: (781)893-8277
;
; INFORMATION FOR SEQ ID NO: 4069:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 92 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; HYPOTHETICAL: YES
; ORIGINAL SOURCE:
; ORGANISM: Enterococcus faecium
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (B) LOCATION 1...92
; SEQUENCE DESCRIPTION: SEQ ID NO: 4069:
;
; US-09-107-532A-4069
;
; Query Match 64.3%; Score 27; DB 4; Length 92;
; Best Local Similarity 62.5%; Pred. No. 1e+02;
; Matches 5; Conservative 1; Mismatches 2; Indels 0; Gaps 0;
;
; QY 1 NVLMAMNI 8
; |.:|.:|.:|
; Db 53 NILFKMNI 60
;
; RESULT 22
; US-08-031-538-68
; Sequence 68, Application US/08031538
; Patent No. 5968817
; GENERAL INFORMATION:
; APPLICANT: Sutcliffe, J Gregor
; APPLICANT: Erlander, Mark G
; APPLICANT: Lovenberg, Timothy W
; TITLE OF INVENTION: HUMAN SEROTONIN RECEPTORS, DNA ENCODING
; TITLE OF INVENTION: THE RECEPTORS, AND USES THEREOF
; NUMBER OF SEQUENCES: 73
; CORRESPONDENCE ADDRESSES:
; ADDRESSEE: The Scripps Research Institute, Office of
; ADDRESSEE: Patent Counsel
; STREET: 10666 No. 5968817th Torrey Pines Road, TPC 8
; CITY: La Jolla
; STATE: CA
; COUNTRY: USA
; ZIP: 92037
;
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/031,538
; FILING DATE: 19930315
; CLASSIFICATION: 800
; ATTORNEY/AGENT INFORMATION:
; NAME: Fitting, Thomas
; REGISTRATION NUMBER: 34,163

```

us-09-900-147-2.ra1

Tue Feb 17 11:55:47 2004

```

REFERENCE/DOCKET NUMBER:  TSR5099P
TELECOMMUNICATION INFORMATION:
TELEPHONE:  619-554-2937
TELEFAX:  619-554-6312
INFORMATION FOR SEQ ID NO:  68:
SEQUENCE CHARACTERISTICS:
LENGTH:  31 amino acids
TYPE:  AMINO ACID
TOPOLOGY:  linear
MOLECULE TYPE:  peptide
FEATURE:
NAME/KEY:  Region
LOCATION:  (7*8)
OTHER INFORMATION:  /note= "An unspecified amino acid
residue is present in this location"
US-08-031-538-68

Query Match      59.5%; Score 25; DB 2; Length 31;
Best Local Similarity  44.4%; Pred. No. 76;
Matches  4; Conservative  4; Mismatches  1; Indels  0; Gaps  0;

Qy      1 NVLMAMNII 9
Db      21 NVFIAMDVM 29

RESULT 23
US-09-134-001C-3106
Sequence 3106, Application US/09134001C
Patent No. 6380370
GENERAL INFORMATION:
APPLICANT:  Lynn Doucette-Stamm et al
TITLE OF INVENTION:  NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO STAPHYLOCOCCUS
EPIDERMIDIS FOR DIAGNOSTICS AND THERAPEUTICS
FILE REFERENCE:  GTC-007
CURRENT APPLICATION NUMBER:  US/09/134,001C
CURRENT FILING DATE:  1998-08-13
PRIOR APPLICATION NUMBER:  US 60/064,964
PRIOR FILING DATE:  1997-11-08
PRIOR APPLICATION NUMBER:  US 60/055,779
PRIOR FILING DATE:  1997-08-14
NUMBER OF SEQ ID NOS:  5674
SEQ ID NO 3106
LENGTH:  62
TYPE:  PRT
ORGANISM:  Staphylococcus epidermidis
US-09-134-001C-3106

Query Match      59.5%; Score 25; DB 4; Length 62;
Best Local Similarity  71.4%; Pred. No. 1.7e+02;
Matches  5; Conservative  1; Mismatches  1; Indels  0; Gaps  0;

Qy      2 VLMAMNI 8
Db      32 VLMTLNI 38

RESULT 24
US-09-205-258-1030
Sequence 1030, Application US/09205258
Patent No. 6525174
GENERAL INFORMATION:
APPLICANT:  Young et al.
TITLE OF INVENTION:  207 Human Secreted Proteins
FILE REFERENCE:  PZ007P1
CURRENT APPLICATION NUMBER:  US/09/205,258
CURRENT FILING DATE:  1998-12-04
EARLIER APPLICATION NUMBER:  PCT/US98/11422
EARLIER FILING DATE:  1998-06-04
EARLIER APPLICATION NUMBER:  60/048,885
EARLIER FILING DATE:  1997-06-06
EARLIER APPLICATION NUMBER:  60/049,375
EARLIER FILING DATE:  1997-06-06
EARLIER APPLICATION NUMBER:  60/048,877
EARLIER FILING DATE:  1997-06-06
EARLIER APPLICATION NUMBER:  60/048,898
EARLIER FILING DATE:  1997-06-06
EARLIER APPLICATION NUMBER:  60/048,962
EARLIER FILING DATE:  1997-06-06
EARLIER APPLICATION NUMBER:  60/048,963
EARLIER FILING DATE:  1997-06-06
EARLIER APPLICATION NUMBER:  60/048,877
EARLIER FILING DATE:  1997-06-06
EARLIER APPLICATION NUMBER:  60/048,878
EARLIER FILING DATE:  1997-06-06
EARLIER APPLICATION NUMBER:  60/070,923
EARLIER FILING DATE:  1997-12-18
EARLIER APPLICATION NUMBER:  60/092,921
EARLIER FILING DATE:  1998-07-15
EARLIER APPLICATION NUMBER:  60/094,657
EARLIER FILING DATE:  1997-06-06
EARLIER APPLICATION NUMBER:  60/048,881
EARLIER FILING DATE:  1997-06-06
EARLIER APPLICATION NUMBER:  60/048,880
EARLIER FILING DATE:  1997-06-06
EARLIER APPLICATION NUMBER:  60/048,896
EARLIER FILING DATE:  1997-06-06
EARLIER APPLICATION NUMBER:  60/049,020
EARLIER FILING DATE:  1997-06-06
EARLIER APPLICATION NUMBER:  60/048,876
EARLIER FILING DATE:  1997-06-06
EARLIER APPLICATION NUMBER:  60/048,895
EARLIER FILING DATE:  1997-06-06
EARLIER APPLICATION NUMBER:  60/048,884
EARLIER FILING DATE:  1997-06-06
EARLIER APPLICATION NUMBER:  60/048,894
EARLIER FILING DATE:  1997-06-06
EARLIER APPLICATION NUMBER:  60/048,971
EARLIER FILING DATE:  1997-06-06
EARLIER APPLICATION NUMBER:  60/048,964
EARLIER FILING DATE:  1997-06-06
EARLIER APPLICATION NUMBER:  60/048,882
EARLIER FILING DATE:  1997-06-06
EARLIER APPLICATION NUMBER:  60/048,899
EARLIER FILING DATE:  1997-06-06
EARLIER APPLICATION NUMBER:  60/048,893
EARLIER FILING DATE:  1997-06-06
EARLIER APPLICATION NUMBER:  60/048,900
EARLIER FILING DATE:  1997-06-06
EARLIER APPLICATION NUMBER:  60/048,901
EARLIER FILING DATE:  1997-06-06
EARLIER APPLICATION NUMBER:  60/048,892
EARLIER FILING DATE:  1997-06-06
EARLIER APPLICATION NUMBER:  60/048,915
EARLIER FILING DATE:  1997-06-06
EARLIER APPLICATION NUMBER:  60/049,019
EARLIER FILING DATE:  1997-06-06
EARLIER APPLICATION NUMBER:  60/048,970
EARLIER FILING DATE:  1997-06-06
EARLIER APPLICATION NUMBER:  60/048,972
EARLIER FILING DATE:  1997-06-06
EARLIER APPLICATION NUMBER:  60/048,916
EARLIER FILING DATE:  1997-06-06
EARLIER APPLICATION NUMBER:  60/049,373
EARLIER FILING DATE:  1997-06-06
EARLIER APPLICATION NUMBER:  60/048,875
EARLIER FILING DATE:  1997-06-06
EARLIER APPLICATION NUMBER:  60/049,374
EARLIER FILING DATE:  1997-06-06
EARLIER APPLICATION NUMBER:  60/048,917
EARLIER FILING DATE:  1997-06-06
EARLIER APPLICATION NUMBER:  60/048,949
EARLIER FILING DATE:  1997-06-06
EARLIER APPLICATION NUMBER:  60/048,974
EARLIER FILING DATE:  1997-06-06
EARLIER APPLICATION NUMBER:  60/048,883
EARLIER FILING DATE:  1997-06-06
EARLIER APPLICATION NUMBER:  60/048,897
EARLIER FILING DATE:  1997-06-06
EARLIER APPLICATION NUMBER:  60/048,898
EARLIER FILING DATE:  1997-06-06
EARLIER APPLICATION NUMBER:  60/048,962
EARLIER FILING DATE:  1997-06-06
EARLIER APPLICATION NUMBER:  60/048,963
EARLIER FILING DATE:  1997-06-06
EARLIER APPLICATION NUMBER:  60/048,877
EARLIER FILING DATE:  1997-06-06
EARLIER APPLICATION NUMBER:  60/048,878
EARLIER FILING DATE:  1997-06-06
EARLIER APPLICATION NUMBER:  60/070,923
EARLIER FILING DATE:  1997-12-18
EARLIER APPLICATION NUMBER:  60/092,921
EARLIER FILING DATE:  1998-07-15
EARLIER APPLICATION NUMBER:  60/094,657
EARLIER FILING DATE:  1997-06-06

```



; EARLIER FILING DATE: 1998-07-30  
; NUMBER OF SEQ ID NOS: 1227  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 1030  
; LENGTH: 66  
; TYPE: PR1  
; ORGANISM: Homo sapiens  
US-09-205-258-1030

Query Match 59.5%; Score 25; DB 4; Length 66;  
Best Local Similarity 44.4%; Pred. No. 1.8e+02;  
Matches 4; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 1 NVLMAMNII 9  
Db 58 NILKALGIV 66

RESULT 25  
US-09-205-258-1032  
; Sequence 1032, Application US/09205258  
; Patent No. 6525174  
; GENERAL INFORMATION:  
; APPLICANT: Young et al.  
; TITLE OF INVENTION: 207 Human Secreted Proteins  
; FILE REFERENCE: P2007P1  
; CURRENT APPLICATION NUMBER: US/09/205,258  
; CURRENT FILING DATE: 1998-12-04  
; EARLIER APPLICATION NUMBER: PCT/US98/11422  
; EARLIER FILING DATE: 1998-06-04  
; EARLIER APPLICATION NUMBER: 60/048,885  
; EARLIER FILING DATE: 1997-06-06  
; EARLIER APPLICATION NUMBER: 60/049,375  
; EARLIER FILING DATE: 1997-06-06  
; EARLIER APPLICATION NUMBER: 60/048,881  
; EARLIER FILING DATE: 1997-06-06  
; EARLIER APPLICATION NUMBER: 60/048,880  
; EARLIER FILING DATE: 1997-06-06  
; EARLIER APPLICATION NUMBER: 60/048,896  
; EARLIER FILING DATE: 1997-06-06  
; EARLIER APPLICATION NUMBER: 60/049,020  
; EARLIER FILING DATE: 1997-06-06  
; EARLIER APPLICATION NUMBER: 60/048,876  
; EARLIER FILING DATE: 1997-06-06  
; EARLIER APPLICATION NUMBER: 60/048,895  
; EARLIER FILING DATE: 1997-06-06  
; EARLIER APPLICATION NUMBER: 60/048,884  
; EARLIER FILING DATE: 1997-06-06  
; EARLIER APPLICATION NUMBER: 60/048,894  
; EARLIER FILING DATE: 1997-06-06  
; EARLIER APPLICATION NUMBER: 60/048,971  
; EARLIER FILING DATE: 1997-06-06  
; EARLIER APPLICATION NUMBER: 60/048,964  
; EARLIER FILING DATE: 1997-06-06  
; EARLIER APPLICATION NUMBER: 60/048,882  
; EARLIER FILING DATE: 1997-06-06  
; EARLIER APPLICATION NUMBER: 60/048,899  
; EARLIER FILING DATE: 1997-06-06  
; EARLIER APPLICATION NUMBER: 60/048,893  
; EARLIER FILING DATE: 1997-06-06  
; EARLIER APPLICATION NUMBER: 60/048,900  
; EARLIER FILING DATE: 1997-06-06  
; EARLIER APPLICATION NUMBER: 60/048,901  
; EARLIER FILING DATE: 1997-06-06  
; EARLIER APPLICATION NUMBER: 60/048,892  
; EARLIER FILING DATE: 1997-06-06  
; EARLIER APPLICATION NUMBER: 60/048,915  
; EARLIER FILING DATE: 1997-06-06  
; EARLIER APPLICATION NUMBER: 60/049,019  
; EARLIER FILING DATE: 1997-06-06  
; EARLIER APPLICATION NUMBER: 60/048,970  
; EARLIER FILING DATE: 1997-06-06  
; EARLIER APPLICATION NUMBER: 60/048,972

; EARLIER FILING DATE: 1997-06-06  
; EARLIER APPLICATION NUMBER: 60/048,916  
; EARLIER FILING DATE: 1997-06-06  
; EARLIER APPLICATION NUMBER: 60/049,373  
; EARLIER FILING DATE: 1997-06-06  
; EARLIER APPLICATION NUMBER: 60/048,875  
; EARLIER FILING DATE: 1997-06-06  
; EARLIER APPLICATION NUMBER: 60/049,374  
; EARLIER FILING DATE: 1997-06-06  
; EARLIER APPLICATION NUMBER: 60/048,917  
; EARLIER FILING DATE: 1997-06-06  
; EARLIER APPLICATION NUMBER: 60/048,949  
; EARLIER FILING DATE: 1997-06-06  
; EARLIER APPLICATION NUMBER: 60/048,974  
; EARLIER FILING DATE: 1997-06-06  
; EARLIER APPLICATION NUMBER: 60/048,883  
; EARLIER FILING DATE: 1997-06-06  
; EARLIER APPLICATION NUMBER: 60/048,897  
; EARLIER FILING DATE: 1997-06-06  
; EARLIER APPLICATION NUMBER: 60/048,898  
; EARLIER FILING DATE: 1997-06-06  
; EARLIER APPLICATION NUMBER: 60/048,962  
; EARLIER FILING DATE: 1997-06-06  
; EARLIER APPLICATION NUMBER: 60/048,963  
; EARLIER FILING DATE: 1997-06-06  
; EARLIER APPLICATION NUMBER: 60/048,877  
; EARLIER FILING DATE: 1997-06-06  
; EARLIER APPLICATION NUMBER: 60/048,878  
; EARLIER FILING DATE: 1997-06-06  
; EARLIER APPLICATION NUMBER: 60/070,923  
; EARLIER FILING DATE: 1997-12-18  
; EARLIER APPLICATION NUMBER: 60/092,921  
; EARLIER FILING DATE: 1998-07-15  
; EARLIER APPLICATION NUMBER: 60/094,657  
; EARLIER FILING DATE: 1998-07-30  
; NUMBER OF SEQ ID NOS: 1227  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 1032  
; LENGTH: 84  
; TYPE: PR1  
; ORGANISM: Homo sapiens  
US-09-205-258-1032

Query Match 59.5%; Score 25; DB 4; Length 84;  
Best Local Similarity 44.4%; Pred. No. 2.3e+02;  
Matches 4; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 1 NVLMAMNII 9  
Db 71 NILKALGIV 79

RESULT 26  
US-09-134-001C-4976  
; Sequence 4976, Application US/09134001C  
; Patent No. 6380370  
; GENERAL INFORMATION:  
; APPLICANT: Lynn Doucette-Stamm et al  
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO STAPHYLOCOCCUS  
; FILE REFERENCE: GTC-007  
; CURRENT APPLICATION NUMBER: US/09/134,001C  
; CURRENT FILING DATE: 1998-08-13  
; PRIOR APPLICATION NUMBER: US 60/064,964  
; PRIOR FILING DATE: 1997-11-08  
; PRIOR APPLICATION NUMBER: US 60/055,779  
; PRIOR FILING DATE: 1997-08-14  
; NUMBER OF SEQ ID NOS: 5674  
; SEQ ID NO 4976  
; LENGTH: 93  
; TYPE: PR1  
; ORGANISM: Staphylococcus epidermidis  
US-09-134-001C-4976

Tue Feb 17 11:55:47 2004

us-09-900-147-2.ra1

```
Query Match          59.5%; Score 25; DB 4; Length 93;
Best Local Similarity 55.8%; Pred. No. 2.6e+02;
Matches 5; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

Qy 1 NVLMANNII 9
Db 27 NVLEVGII 35

RESULT 27
US-09-328-352-8011
; Sequence 8011, Application US/09328352
; Patent No. 6562958
; GENERAL INFORMATION:
; APPLICANT: Gary L. Breton et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO ACINETOBACTER
; TITLE OF INVENTION: BAUMANNII FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: GTC99-03PA
; CURRENT APPLICATION NUMBER: US/09/328,352
; CURRENT FILING DATE: 1999-06-04
; NUMBER OF SEQ ID NOS: 8252
; SEQ ID NO 8011
; LENGTH: 95
; TYPE: PRT
; ORGANISM: Acinetobacter baumannii
US-09-328-352-8011

Query Match          59.5%; Score 25; DB 4; Length 95;
Best Local Similarity 50.0%; Pred. No. 2.7e+02;
Matches 4; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

Qy 2 VLMMANNII 9
Db 9 IVMAANVI 16

RESULT 28
US-07-626-618A-6
; Sequence 6, Application US/07626618A
; Patent No. 5422265
; GENERAL INFORMATION:
; APPLICANT: Van Tol, Hubert H.M.
; APPLICANT: Civelli, Olivier
; TITLE OF INVENTION: A No. 5422265el Human Dopamine Receptor and Uses
; NUMBER OF SEQUENCES: 22
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Allegretti & Witcoff, Ltd.
; STREET: 10 South Wacker Drive, Suite 3000
; CITY: Chicago
; STATE: Illinois
; COUNTRY: USA
; ZIP: 60606
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; FILING DATE: 7 DEC 1990
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: No. 5422265nan, Kevin E
; REGISTRATION NUMBER: 35,303
; REFERENCE/DOCKET NUMBER: 90,1092
; TELEPHONE: 312-715-1000
; TELEFAX: 312-715-1234
; TELEX: 810-221-8317
; INFORMATION FOR SEQ ID NO: 6:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 37 amino acids
; TYPE: AMINO ACID
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-07-626-618A-6

Query Match          57.1%; Score 24; DB 1; Length 37;
Best Local Similarity 44.4%; Pred. No. 1.5e+02;
Matches 4; Conservative 4; Mismatches 1; Indels 0; Gaps 0;

Qy 1 NVLMANNII 9
Db 14 DALMAMDVM 22

RESULT 30
US-08-333-977-6
; Sequence 6, Application US/08333977
; Patent No. 5594108
; GENERAL INFORMATION:
; APPLICANT: Van Tol, Hubert H.M.
; APPLICANT: Civelli, Olivier
; TITLE OF INVENTION: A No. 5594108el Human Dopamine Receptor and Uses
; NUMBER OF SEQUENCES: 22
```

;; CORRESPONDENCE ADDRESS:  
;; ADDRESSEE: Allegretti & Witcoff, Ltd.  
;; STREET: 10 South Wacker Drive, Suite 3000  
;; CITY: Chicago  
;; STATE: Illinois  
;; COUNTRY: USA  
;; ZIP: 60606  
;;  
;; COMPUTER READABLE FORM:  
;; MEDIUM TYPE: Floppy disk  
;; COMPUTER: IBM PC compatible  
;; OPERATING SYSTEM: PC-DOS/MS-DOS  
;; SOFTWARE: PatentIn Release #1.0, Version #1.25  
;; CURRENT APPLICATION DATA: US/08/333,977  
;; APPLICATION NUMBER: US/08/333,977  
;; FILING DATE: 03-NOV-1994  
;; CLASSIFICATION: 530  
;; PRIOR APPLICATION DATA:  
;; APPLICATION NUMBER: US 07/626,618  
;; FILING DATE: 7 DEC 1990  
;; ATTORNEY/AGENT INFORMATION:  
;; NAME: No. 5594108nan, Kevin E  
;; REGISTRATION NUMBER: 35,303  
;; REFERENCE/DOCKET NUMBER: 90,1092  
;; TELECOMMUNICATION INFORMATION:  
;; TELEPHONE: 312-715-1000  
;; TELEFAX: 312-715-1234  
;; TELEX: 810-221-8317  
;; INFORMATION FOR SEQ ID NO: 6:  
;; SEQUENCE CHARACTERISTICS:  
;; LENGTH: 37 amino acids  
;; TYPE: amino acid  
;; TOPOLOGY: linear  
;; MOLECULE TYPE: protein  
US-08-333-977-6

Query Match 57.1%; Score 24; DB 1; Length 37;  
Best Local Similarity 44.4%; Pred. No. 1.5e+02;  
Matches 4; Conservative 4; Mismatches 1; Indels 0; Gaps 0;

QY 1 NVLMAMNII 9  
Db 14 DALMAMDMV 22

RESULT 31  
US-08-487-811A-6  
; Sequence 6, Application US/08487811A  
; Patent No. 5883226  
; GENERAL INFORMATION:  
; APPLICANT: Civelli, Olivier  
; APPLICANT: Van Tol, Hubert H.M.  
; TITLE OF INVENTION: A No. 5883226el Human Dopamine Receptor and Uses  
; NUMBER OF SEQUENCES: 24  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: McDonnell Boehnen Hulbert & Berghoff  
; STREET: 300 South Wacker Drive  
; CITY: Chicago  
; STATE: IL  
; COUNTRY: USA  
; ZIP: 60606  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/487,811A  
; FILING DATE: 07-JUN-1995  
; CLASSIFICATION: 530  
; ATTORNEY/AGENT INFORMATION:  
; NAME: No. 5883226nan, Kevin E  
; REGISTRATION NUMBER: 35,303  
; REFERENCE/DOCKET NUMBER: 90,1092-L

;; TELECOMMUNICATION INFORMATION:  
;; TELEPHONE: 312-913-0001  
;; TELEFAX: 312-913-0002  
;; TELEX:  
;; INFORMATION FOR SEQ ID NO: 6:  
;; SEQUENCE CHARACTERISTICS:  
;; LENGTH: 37 amino acids  
;; TYPE: amino acid  
;; TOPOLOGY: linear  
;; MOLECULE TYPE: protein  
US-08-487-811A-6

Query Match 57.1%; Score 24; DB 2; Length 37;  
Best Local Similarity 44.4%; Pred. No. 1.5e+02;  
Matches 4; Conservative 4; Mismatches 1; Indels 0; Gaps 0;

QY 1 NVLMAMNII 9  
Db 14 DALMAMDMV 22

RESULT 32  
US-09-060-694-6  
; Sequence 6, Application US/09060694  
; Patent No. 6203998  
; GENERAL INFORMATION:  
; APPLICANT: Civelli, Olivier  
; APPLICANT: Van Tol, Hubert H.M.  
; TITLE OF INVENTION: A No. 6203998el Human Dopamine Receptor and Uses  
; NUMBER OF SEQUENCES: 24  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: McDonnell Boehnen Hulbert & Berghoff  
; STREET: 300 South Wacker Drive  
; CITY: Chicago  
; STATE: IL  
; COUNTRY: USA  
; ZIP: 60606  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/060,694  
; FILING DATE: 15-APR-1998  
; CLASSIFICATION: 530  
; ATTORNEY/AGENT INFORMATION:  
; NAME: No. 6203998nan, Kevin E  
; REGISTRATION NUMBER: 35,303  
; REFERENCE/DOCKET NUMBER: 90,1092-MM  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 312-913-0001  
; TELEFAX: 312-913-0002  
; TELEX:  
; INFORMATION FOR SEQ ID NO: 6:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 37 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
US-09-060-694-6

Query Match 57.1%; Score 24; DB 3; Length 37;  
Best Local Similarity 44.4%; Pred. No. 1.5e+02;  
Matches 4; Conservative 4; Mismatches 1; Indels 0; Gaps 0;

QY 1 NVLMAMNII 9  
Db 14 DALMAMDMV 22

RESULT 33  
US-09-378-074-6

us-09-900-147-2.ra1

Tue Feb 17 11:55:47 2004

```

; Sequence 6, Application US/09378074
; Patent No. 6437114
; GENERAL INFORMATION:
; APPLICANT: Van Tol, Hubert H.M.
; TITLE OF INVENTION: A Novel Human Dopamine Receptor and Uses
; NUMBER OF SEQUENCES: 22
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Allegretti & Witcoff, Ltd. 3000
; STREET: 10 South Wacker Drive, Suite 3000
; CITY: Chicago
; STATE: Illinois
; COUNTRY: USA
; ZIP: 60606
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent in Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/378,074
; FILING DATE: 20-Aug-1999
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 07/928,611
; FILING DATE: <Unknown>
; ATTORNEY/AGENT INFORMATION:
; NAME: No. 6437114nan, Kevin E
; REGISTRATION NUMBER: 35,303
; REFERENCE/DOCKET NUMBER: 90,1092-B
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 312-715-1000
; TELEFAX: 312-715-1234
; TELEX: 810-221-8317
; INFORMATION FOR SEQ ID NO: 6:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 37 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; SEQUENCE DESCRIPTION: SEQ ID NO: 6:
US-09-378-074-6

Query Match 57.1%; Score 24; DB 4; Length 37;
Best Local Similarity 44.4%; Pred. No. 1.5e+02;
Matches 4; Conservative 4; Mismatches 1; Indels 0; Gaps 0;

QY 1 NVLMAMNII 9
Db 14 DALMAMDVM 22

RESULT 34
PCT-US93-07370-6
; Sequence 6, Application PC/TUS9307370
; GENERAL INFORMATION:
; APPLICANT:
; TITLE OF INVENTION: A Novel Human Dopamine Receptor and Uses
; NUMBER OF SEQUENCES: 22
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent in Release #1.0, Version #1.25 (BPO)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US93/07370
; INFORMATION FOR SEQ ID NO: 6:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 37 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; PCT-US93-07370-6

```

```

Query Match 57.1%; Score 24; DB 5; Length 37;
Best Local Similarity 44.4%; Pred. No. 1.5e+02;
Matches 4; Conservative 4; Mismatches 1; Indels 0; Gaps 0;

QY 1 NVLMAMNII 9
Db 14 DALMAMDVM 22

RESULT 35
US-09-215-221-18
; Sequence 18, Application US/09215221
; Patent No. 6265562
; GENERAL INFORMATION:
; APPLICANT: EILERS, MARTIN
; APPLICANT: BUERGIN, ANDREA
; APPLICANT: SEDLACEK, HANS-HARALD
; TITLE OF INVENTION: NUCLEIC ACID CONSTRUCTS WHOSE ACTIVITY IS AFFECTED BY
; TITLE OF INVENTION: INHIBITORS OF CYCLIN-DEPENDANT KINASES AND USES THEREOF
; FILE REFERENCE: 026083/0192
; CURRENT APPLICATION NUMBER: US/09/215,221
; CURRENT FILING DATE: 1998-12-18
; PRIOR APPLICATION NUMBER: 197 56 975.7
; PRIOR FILING DATE: 1997-12-20
; NUMBER OF SEQ ID NOS: 57
; SOFTWARE: Patent In Ver. 2.1
; SEQ ID NO 18
; LENGTH: 40
; TYPE: PRT
; ORGANISM: Marine sp.
US-09-215-221-18

Query Match 57.1%; Score 24; DB 3; Length 40;
Best Local Similarity 62.5%; Pred. No. 1.6e+02;
Matches 5; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 NVLMAMNI 8
Db 7 NVLIAPNM 14

RESULT 36
US-09-252-991A-30061
; Sequence 30061, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 30061
; LENGTH: 69
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-30061

Query Match 57.1%; Score 24; DB 4; Length 69;
Best Local Similarity 62.5%; Pred. No. 3e+02;
Matches 5; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 NVLMAMNI 8
Db 37 SVLMALKI 44

```

```
RESULT 37
US-09-134-001C-4499
; Sequence 4499, Application US/09134001C
; Patent No. 6380370
; GENERAL INFORMATION:
; APPLICANT: Lynn Doucette-Stamm et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO STAPHYLOCOCCUS
; FILE REFERENCE: EPIDERMIDIS FOR DIAGNOSTICS AND THERAPEUTICS
; CURRENT APPLICATION NUMBER: US/09/134.001C
; PRIOR FILING DATE: 1998-08-13
; PRIOR APPLICATION NUMBER: US 60/064,964
; PRIOR FILING DATE: 1997-11-08
; PRIOR APPLICATION NUMBER: US 60/055,779
; PRIOR FILING DATE: 1997-08-14
; NUMBER OF SEQ ID NOS: 5674
; SEQ ID NO 4499
; LENGTH: 76
; TYPE: PRT
; ORGANISM: Staphylococcus epidermidis
US-09-134-001C-4499

Query Match 57.1%; Score 24; DB 4; Length 76;
Best Local Similarity 57.1%; Pred. No. 3.4e+02;
Matches 4; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 3 LMAMNII 9
   :||:|
Db 30 ILALNII 36

RESULT 38
US-09-328-352-5816
; Sequence 5816, Application US/09328352
; Patent No. 6562958
; GENERAL INFORMATION:
; APPLICANT: Gary L. Breton et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO ACINETOBACTER
; FILE REFERENCE: GTC99-03PA
; CURRENT APPLICATION NUMBER: US/09/328.352
; CURRENT FILING DATE: 1999-06-04
; NUMBER OF SEQ ID NOS: 8252
; SEQ ID NO 5816
; LENGTH: 77
; TYPE: PRT
; ORGANISM: Acinetobacter baumannii
US-09-328-352-5816

Query Match 57.1%; Score 24; DB 4; Length 77;
Best Local Similarity 50.0%; Pred. No. 3.4e+02;
Matches 4; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 1 NVLMAMNII 8
   |::|||
Db 1 NLVLKNNI 8

RESULT 39
US-09-103-478-27
; Sequence 27, Application US/09103478
; Patent No. 6235975
; GENERAL INFORMATION:
; APPLICANT: Harada, John
; APPLICANT: Lotan, Tamar
; APPLICANT: Ohto, Masa-aki
; APPLICANT: Goldberg, Robert B.
; APPLICANT: Fischer, Robert L.
; TITLE OF INVENTION: LEAFY COTYLEDON1 Genes and Their Uses
; NUMBER OF SEQUENCES: 29
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Townsend and Crew LLP
; STREET: Two Embarcadero Center, Eighth Floor
```

```
; CITY: San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94111-3834
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/103,478
; FILING DATE: 24-JUN-1998
; CLASSIFICATION: 800
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 09/026,221
; FILING DATE: 19-FEB-1998
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/804,534
; FILING DATE: 21-FEB-1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Einhorn, Gregory P.
; REGISTRATION NUMBER: 38,440
; REFERENCE/DOCKET NUMBER: 023070-077611US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 576-0200
; TELEFAX: (415) 576-0300
; INFORMATION FOR SEQ ID NO: 27:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 85 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-09-103-478-27

Query Match 57.1%; Score 24; DB 3; Length 85;
Best Local Similarity 44.4%; Pred. No. 3.8e+02;
Matches 4; Conservative 4; Mismatches 1; Indels 0; Gaps 0;

QY 1 NVLMAMNII 9
   :||:|
Db 60 DVLLALNTL 68

RESULT 40
US-09-193-931C-27
; Sequence 27, Application US/09193931C
; Patent No. 6320102
; GENERAL INFORMATION:
; APPLICANT: Harada, John
; APPLICANT: Lotan, Tamar
; APPLICANT: Ohto, Masa-aki
; APPLICANT: Goldberg, Robert B.
; APPLICANT: Fischer, Robert L.
; APPLICANT: The Regents of the University of California
; TITLE OF INVENTION: LEAFY COTYLEDON1 Genes and Their Uses
; FILE REFERENCE: 023070-077620
; CURRENT APPLICATION NUMBER: US/09/193,931C
; CURRENT FILING DATE: 1998-11-17
; PRIOR APPLICATION NUMBER: US 09/103,478
; PRIOR FILING DATE: 1998-06-24
; PRIOR APPLICATION NUMBER: US 09/026,221
; PRIOR FILING DATE: 1998-02-19
; PRIOR APPLICATION NUMBER: US 08/804,534
; PRIOR FILING DATE: 1997-02-21
; NUMBER OF SEQ ID NOS: 29
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 27
; LENGTH: 85
; TYPE: PRT
; ORGANISM: Schizosaccharomyces pombe
; FEATURE:
; NAME/KEY: DOMAIN
```

Tue Feb 17 11:55:47 2004

us-09-900-147-2.ra1

LOCATION: (1)..(85)  
OTHER INFORMATION: S. pombe HAP3 subunit of CCAAT box-binding factor  
OTHER INFORMATION: (CBF) protein B domain homolog  
US-09-193-931C-27

Query Match 57.1%; Score 24; DB 4; Length 85;  
Best Local Similarity 44.4%; Pred. No. 3.9e+02;  
Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 N1VLMAMNII 9  
Db 60 DVLALNTL 68

RESULT 41  
US-08-289-247B-4  
Sequence 4, Application US/08289247B  
Patent No. 5728579  
GENERAL INFORMATION:  
APPLICANT: Morrison, Briggs W.

TITLE OF INVENTION: Detection and Treatment of Breast  
APPLICANT: Leder, Philip  
TITLE OF INVENTION: Cancer  
NUMBER OF SEQUENCES: 4  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Clark & Elbing LLP  
STREET: 176 Federal Street  
CITY: Boston  
STATE: Massachusetts  
COUNTRY: U.S.A.  
ZIP: 02110-2223

COMPUTER READABLE FORM:  
MEDIUM TYPE: 3.5" Diskette, 1.44 Mb  
COMPUTER: IBM PS/2 Model 502 or 55SX  
OPERATING SYSTEM: MS-DOS (Version 5.0)  
SOFTWARE: WordPerfect (Version 5.1)  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/289,247B  
FILING DATE: August 11, 1994  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER:  
FILING DATE:

ATTORNEY/AGENT INFORMATION:  
NAME: Bieker-Brady, Kristina  
REGISTRATION NUMBER: 39,109  
REFERENCE/DOCKET NUMBER: 00383/021001  
TELEPHONE: (617) 428-0200  
TELEFAX: (617) 428-7045  
TELEX:

INFORMATION FOR SEQ ID NO: 4:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 87  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
US-08-289-247B-4

Query Match 57.1%; Score 24; DB 1; Length 87;  
Best Local Similarity 75.0%; Pred. No. 3.9e+02;  
Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 2 V1MAMNII 9  
Db 47 VLCAMGII 54

RESULT 42  
US-08-725-531-4  
Sequence 4, Application US/08725531  
Patent No. 5756310  
GENERAL INFORMATION:  
APPLICANT: Bandman, Olga

APPLICANT: Goli, Surya K.  
TITLE OF INVENTION: NOVEL HUMAN PHOSPHOLEMMAN-LIKE PROTEIN  
NUMBER OF SEQUENCES: 6  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Incyte Pharmaceuticals, Inc.

STREET: 3174 Porter Drive  
CITY: Palo Alto  
STATE: CA  
COUNTRY: U.S.  
ZIP: 94304

COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette  
COMPUTER: IBM Compatible  
OPERATING SYSTEM: DOS  
SOFTWARE: FastSeq Version 1.5  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/725,531  
FILING DATE: Filed Herewith  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER:

FILING DATE:  
ATTORNEY/AGENT INFORMATION:  
NAME: Billings, Lucy J.  
REGISTRATION NUMBER: 36,749  
REFERENCE/DOCKET NUMBER: PF-0128 US  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 415-855-0555  
TELEFAX: 415-845-4166

INFORMATION FOR SEQ ID NO: 4:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 87 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
IMMEDIATE SOURCE:  
LIBRARY: GenBank  
CLONE: 1085026  
US-08-725-531-4

Query Match 57.1%; Score 24; DB 1; Length 87;  
Best Local Similarity 75.0%; Pred. No. 3.9e+02;  
Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 2 V1MAMNII 9  
Db 47 VLCAMGII 54

RESULT 43  
US-08-738-127-4  
Sequence 4, Application US/08738127  
Patent No. 5919655  
GENERAL INFORMATION:  
APPLICANT: Bandman, Olga  
APPLICANT: Hawkins, Phillip R.  
TITLE OF INVENTION: NOVEL HUMAN PHOSPHOLEMMAN HOMOLOG  
NUMBER OF SEQUENCES: 6  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Incyte Pharmaceuticals, Inc.  
STREET: 3174 Porter Drive  
CITY: Palo Alto  
STATE: CA  
COUNTRY: US  
ZIP: 94304

COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette  
COMPUTER: IBM Compatible  
OPERATING SYSTEM: DOS  
SOFTWARE: FastSeq Version 1.5  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/738,127  
FILING DATE: Filed Herewith

```
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Billings, Lucy J.
; REGISTRATION NUMBER: 36,749
; REFERENCE/DOCKET NUMBER: PF-0141 US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415-855-0555
; TELEFAX: 415-845-4166
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 87 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; IMMEDIATE SOURCE:
; LIBRARY: GenBank
; CLONE: 1085026
; US-08-738-127-4

Query Match 57.1%; Score 24; DB 2; Length 87;
Best Local Similarity 75.0%; Pred. No. 3.9e+02;
Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 2 VLMAMNII 9
Db 47 VLCAMGII 54

RESULT 44
US-09-213-392-4
; Sequence 4, Application US/09213392
; Patent No. 5945505
; GENERAL INFORMATION:
; APPLICANT: Bandman, Olga
; APPLICANT: Goli, Surya K.
; TITLE OF INVENTION: NOVEL HUMAN PHOSPHOLEMMAN-LIKE PROTEIN
; NUMBER OF SEQUENCES: 6
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Incyte Pharmaceuticals, Inc.
; STREET: 3174 Porter Drive
; CITY: Palo Alto
; STATE: CA
; COUNTRY: U.S.
; ZIP: 94304
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq Version 1.5
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/213,392
; FILING DATE:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 09/083,661
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Billings, Lucy J.
; REGISTRATION NUMBER: 36,749
; REFERENCE/DOCKET NUMBER: PF-0128 US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415-855-0555
; TELEFAX: 415-845-4166
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 87 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; IMMEDIATE SOURCE:
; LIBRARY: GenBank
; CLONE: 1085026
; US-09-213-392-4

Query Match 57.1%; Score 24; DB 2; Length 87;
Best Local Similarity 75.0%; Pred. No. 3.9e+02;
Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 2 VLMAMNII 9
Db 47 VLCAMGII 54

RESULT 45
US-09-083-661-4
; Sequence 4, Application US/09083661
; Patent No. 5955283
; GENERAL INFORMATION:
; APPLICANT: Bandman, Olga
; APPLICANT: Goli, Surya K.
; TITLE OF INVENTION: NOVEL HUMAN PHOSPHOLEMMAN-LIKE PROTEIN
; NUMBER OF SEQUENCES: 6
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Incyte Pharmaceuticals, Inc.
; STREET: 3174 Porter Drive
; CITY: Palo Alto
; STATE: CA
; COUNTRY: U.S.
; ZIP: 94304
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq Version 1.5
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/083,661
; FILING DATE:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/725,531
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Billings, Lucy J.
; REGISTRATION NUMBER: 36,749
; REFERENCE/DOCKET NUMBER: PF-0128 US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415-855-0555
; TELEFAX: 415-845-4166
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 87 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; IMMEDIATE SOURCE:
; LIBRARY: GenBank
; CLONE: 1085026
; US-09-083-661-4

Query Match 57.1%; Score 24; DB 2; Length 87;
Best Local Similarity 75.0%; Pred. No. 3.9e+02;
Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 2 VLMAMNII 9
Db 47 VLCAMGII 54

Search completed: February 17, 2004, 10:59:39
Job time : 6.50495 secs
```

```
; LIBRARY: GenBank
; CLONE: 1085026
; US-09-213-392-4

Query Match 57.1%; Score 24; DB 2; Length 87;
Best Local Similarity 75.0%; Pred. No. 3.9e+02;
Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 2 VLMAMNII 9
Db 47 VLCAMGII 54

RESULT 45
US-09-083-661-4
; Sequence 4, Application US/09083661
; Patent No. 5955283
; GENERAL INFORMATION:
; APPLICANT: Bandman, Olga
; APPLICANT: Goli, Surya K.
; TITLE OF INVENTION: NOVEL HUMAN PHOSPHOLEMMAN-LIKE PROTEIN
; NUMBER OF SEQUENCES: 6
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Incyte Pharmaceuticals, Inc.
; STREET: 3174 Porter Drive
; CITY: Palo Alto
; STATE: CA
; COUNTRY: U.S.
; ZIP: 94304
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq Version 1.5
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/083,661
; FILING DATE:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/725,531
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Billings, Lucy J.
; REGISTRATION NUMBER: 36,749
; REFERENCE/DOCKET NUMBER: PF-0128 US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415-855-0555
; TELEFAX: 415-845-4166
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 87 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; IMMEDIATE SOURCE:
; LIBRARY: GenBank
; CLONE: 1085026
; US-09-083-661-4

Query Match 57.1%; Score 24; DB 2; Length 87;
Best Local Similarity 75.0%; Pred. No. 3.9e+02;
Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 2 VLMAMNII 9
Db 47 VLCAMGII 54

Search completed: February 17, 2004, 10:59:39
Job time : 6.50495 secs
```

**THIS PAGE BLANK (USPTO)**



GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: February 17, 2004, 10:50:12 ; Search time 12.7426 Seconds  
(without alignments)  
182.261 Million cell updates/sec

Title: US-09-900-147-2  
Perfect score: 42  
Sequence: 1 NVLMANNII 9

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 830525 seqs, 258052604 residues

Total number of hits satisfying chosen parameters: 146963

Minimum DB seq length: 0  
Maximum DB seq length: 100

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database :

SPTREMBL 23:\*

- 1: sp\_archaea:\*
- 2: sp\_bacteria:\*
- 3: sp\_fungi:\*
- 4: sp\_human:\*
- 5: sp\_invertebrate:\*
- 6: sp\_mammal:\*
- 7: sp\_mhc:\*
- 8: sp\_organelle:\*
- 9: sp\_phase:\*
- 10: sp\_plant:\*
- 11: sp\_rodent:\*
- 12: sp\_virus:\*
- 13: sp\_vertebrate:\*
- 14: sp\_unclassified:\*
- 15: sp\_rvirus:\*
- 16: sp\_bacteriaph:\*
- 17: sp\_archaeap:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match %	Length	ID	Description
1	30	71.4	83	16	Q8EIV8
2	29	69.0	61	16	Q8NX27
3	29	69.0	64	16	Q8NWS0
4	29	69.0	64	16	Q8NWS0
5	29	69.0	70	16	Q99Q85
6	29	69.0	89	16	Q8CP32
7	29	69.0	98	16	Q99U94
8	28	66.7	71	8	Q9XNQ4
9	28	66.7	98	2	Q8VSS7
10	28	66.7	98	8	O21757
11	28	66.7	98	8	O21705
12	28	66.7	98	8	O21712
13	28	66.7	98	8	O21846
14	27	64.3	47	16	Q8CPD5
15	27	64.3	81	16	Q8X8C1
16	27	64.3	81	16	Q8FBE9
					Q8EIV8 shewanella
					Q8nx27 staphylococ
					Q8nws0 staphylococ
					Q8nws0 staphylococ
					Q99q85 staphylococ
					Q8cp32 staphylococ
					Q99u94 staphylococ
					Q9xnd4 boophilus m
					Q8vss7 proteus mir
					O21757 pygathrix b
					O21705 pygathrix b
					O21712 pygathrix r
					O21846 pygathrix r
					Q8cpd5 staphylococ
					Q8x8c1 escherichia
					Q8fbe9 escherichia

17	27	64.3	99	13	Q8AWZ8	Q8awz8 xenopus lae
18	27	64.3	99	17	Q97YM7	Q97ym7 sulfolobus
19	26	61.9	40	16	Q8CQI6	Q8cq16 staphylococ
20	26	61.9	50	8	Q9GCA4	Q9gca4 samia cynth
21	26	61.9	53	16	Q8X3P6	Q8x3f6 escherichia
22	26	61.9	73	16	Q8Z697	Q8z697 salmonella
23	26	61.9	84	8	Q8W977	Q8w977 leptasteria
24	26	61.9	92	16	Q8ZP30	Q8zp30 salmonella
25	26	61.9	95	4	Q96DE8	Q96de8 homo sapien
26	26	61.9	98	8	Q8LWU8	Q8lww8 macrosclid
27	26	61.9	98	16	Q8YG45	Q8yg45 brucella me
28	26	61.9	98	16	Q8G1S9	Q8gl59 brucella su
29	26	61.9	99	8	Q9XNR1	Q9xnr1 boophilus m
30	25	59.5	21	8	Q9MIJ8	Q9mij8 ciccindela w
31	25	59.5	21	8	Q9MIJ9	Q9mij9 ciccindela c
32	25	59.5	21	8	Q9MIJ2	Q9mij2 ciccindela s
33	25	59.5	21	8	Q9MIJ5	Q9mij5 ciccindela m
34	25	59.5	21	8	Q9MIJ6	Q9mij6 ciccindela b
35	25	59.5	21	8	Q9MIJ4	Q9mij4 ciccindela h
36	25	59.5	21	8	Q9MIK0	Q9mik0 ciccindela g
37	25	59.5	24	11	Q88717	Q88717 mus musculu
38	25	59.5	45	5	Q26413	Q26413 cryptospori
39	25	59.5	59	16	Q9K8Y3	Q9k8y3 bacillus ha
40	25	59.5	74	16	Q99VZ5	Q99vz5 staphylococ
41	25	59.5	75	17	Q8UJ33	Q8uzj3 pyrococcus
42	25	59.5	79	16	Q97P38	Q97p38 streptococ
43	25	59.5	81	8	Q8HNB1	Q8hnb1 daktulospha
44	25	59.5	81	8	Q8HNB0	Q8hnb0 daktulospha
45	25	59.5	81	8	Q8HNA9	Q8hna9 daktulospha

#### ALIGNMENTS

#### RESULT 1

Q8EIV8 PRELIMINARY; PRT; 83 AA.

AC Q8EIV8;  
DT 01-MAR-2003 (TrEMBLrel. 23, Created)  
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)  
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)  
DE Conserved hypothetical protein.  
GN S00721.  
OS Shewanella oneidensis.  
OC Bacteria; Proteobacteria; Gammaproteobacteria; Alteromonadales;  
OC Alteromonadaceae; Shewanella.  
OX NCBI\_TaxID=70863;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=MR-1;  
RX MEDLINE=22297686; PubMed=12368813;  
RA Heidelberg J.F., Paulsen I.T., Nelson K.E., Gaidos E.J., Nelson W.C.,  
RA Read T.D., Eisen J.A., Seshadri R., Ward N., Methe B., Clayton R.A.,  
RA Meyer T., Tsapin A., Scott J., Beanan M., Brinkac L., Daugherty S.,  
RA DeBoy R.T., Dodson R.J., Durkin A.S., Haft D.H., Kolonay J.F.,  
RA Madupu R., Peterson J.D., Umayam L.A., White O., Wolf A.M.,  
RA Vamathevan J., Weidman J., Imprim M., Lee K., Berry K., Lee C.,  
RA Mueller J., Khouri H., Gill J., Utterback T.R., McDonald L.A.,  
RA Feldblyum T.V., Smith H.O., Venter J.C., Nealeson K.H., Fraser C.M.,  
RT "Genome sequence of the dissimilatory metal ion-reducing bacterium  
Shewanella oneidensis.";  
RL Nat. Biotechnol. 20:1118-1123 (2002).  
DR EMBL; AE015517; AAN53799.1; -  
DR TIGR; S00721; -  
KW Hypothetical protein; Complete proteome.  
SQ SEQUENCE 83 AA; 9075 MW; ACS08F38ACB345C CRC64;

Query Match 71.4%; Score 30; DB 16; Length 83;  
Best Local Similarity 77.8%; Pred. No. 80;  
Matches 7; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 NVLMANNII 9

|||||

RESULT 4  
Q8NWS0

## RESULT 6

Q8CP32 PRELIMINARY; PRT; 89 AA.  
 AC Q8CP32;  
 DT 01-MAR-2003 (TrEMBLrel. 23, Created)  
 DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)  
 DE 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)  
 DE Truncated transposase.  
 GN SE1219.  
 OS Staphylococcus epidermidis.  
 OC Bacteria; Firmicutes; Bacillales; Staphylococcus.  
 OX NCBI\_TaxID=1282;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=ATCC 12228;  
 RA Zhang Y., Ren S., Li H., Fu G., Lu L., Lu G., Jia J., Tu Y., Qin Z.,  
 RA Chen Z., Wen Y.;  
 RL Submitted (NOV-2002) to the EMBL/GenBank/DBJ databases.  
 RL EMBL; AB016748; AAC04818.1; -;  
 KW Complete proteome.  
 SQ SEQUENCE 89 AA; 10671 MW; 32E07301FE24162D CRC64;

Query Match 69.0%; Score 29; DB 16; Length 89;  
 Best Local Similarity 85.7%; Pred. No. 1.4e+02;  
 Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 2 VLMAMNI 8  
 |||:|

Db 59 VLMALNI 65  
 |||:|

## RESULT 7

Q99U94 PRELIMINARY; PRT; 98 AA.  
 AC Q99U94;  
 DT 01-JUN-2001 (TrEMBLrel. 17, Created)  
 DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)  
 DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)  
 DE Truncated transposase.  
 GN TRUNCATED-SA OR SAV1390 OR SA1222.  
 OS Staphylococcus aureus (strain Mu50 / ATCC 700699), and  
 OS Staphylococcus aureus (strain N315).  
 OC Bacteria; Firmicutes; Bacillales; Staphylococcus.  
 OX NCBI\_TaxID=158878, 158879;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC SPECIES=S. aureus (strain Mu50), and S. aureus (strain N315);  
 RX MEDLINE=21311952; PubMed=11418146;  
 RA Kuroda M., Ohta T., Uchiyama I., Baba T., Yuzawa H., Kobayashi I.,  
 RA Cui L., Oguchi A., Aoki K.-I., Nagai Y., Lian J.-Q., Ito T.,  
 RA Kanamori M., Matsumaru H., Maruyama A., Murakami H., Hosoyama A.,  
 RA Mizutani-Ui Y., Takahashi N.K., Sawano T., Inoue R.-I., Kaito C.,  
 RA Sekinizu K., Hirakawa H., Kuhara S., Goto S., Yabuzaki J.,  
 RA Kanehisa M., Yamashita A., Oshima K., Furuya K., Yoshino C., Shiba T.,  
 RA Hattori M., Ogasawara N., Hayashi H., Hiramatsu K.;  
 RT "Whole genome sequencing of methicillin-resistant Staphylococcus  
 aureus";  
 RL Lancet 357:1225-1240 (2001).  
 DR ENBL; AP003362; BAB57552.1; -;  
 DR ENBL; AP003133; BAB42482.1; -;  
 DR InterPro; IPR002559; Transposase\_11.  
 DR Pfam; PF01609; Transposase\_11; 1.  
 KW Complete proteome.  
 SQ SEQUENCE 98 AA; 11619 MW; 4853E89EB4C8EF17 CRC64;

Query Match 69.0%; Score 29; DB 16; Length 98;  
 Best Local Similarity 85.7%; Pred. No. 1.5e+02;  
 Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 2 VLMAMNI 8  
 |||:|

Db 59 VLMALNI 65  
 |||:|

## RESULT 8

Q9XNQ4 PRELIMINARY; PRT; 71 AA.  
 ID Q9XNQ4  
 AC Q9XNQ4;  
 DT 01-NOV-1999 (TrEMBLrel. 12, Created)  
 DT 01-NOV-1999 (TrEMBLrel. 12, Last sequence update)  
 DT 01-NOV-1999 (TrEMBLrel. 12, Last annotation update)  
 DE NADH dehydrogenase subunit 2 (Fragment).  
 GN ND2.  
 OS Boophilus microplus (Cattle tick).  
 OC Mitochondrion.  
 OC Eukaryota; Metazoa; Arthropoda; Chelicerata; Arachnida; Acari;  
 OC Parasitiformes; Ixodida; Ixodidae; Boophilus.  
 OX NCBI\_TaxID=6941;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=N.  
 RX MEDLINE=99297341; PubMed=10368952;  
 RA Campbell N.J.H., Barker S.C.;  
 RT "The novel mitochondrial gene arrangement of the cattle tick,  
 RT Boophilus microplus: fivefold tandem repetition of a coding region.";  
 RL Mol. Biol. Evol. 16:732-740 (1999).  
 DR EMBL; AF110613; AAD28382.1; -;  
 KW Mitochondrion.  
 FT NON\_TER 1  
 SQ SEQUENCE 71 AA; 8854 MW; 2FFAB88E6B29B7B2 CRC64;

Query Match 66.7%; Score 28; DB 8; Length 71;  
 Best Local Similarity 55.8%; Pred. No. 1.9e+02;  
 Matches 5; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 1 NVLMAMNII 9  
 |:-|:-|

Db 53 NIYMNMLI 61  
 |:-|:-|

## RESULT 9

Q8VS57 PRELIMINARY; PRT; 98 AA.  
 AC Q8VS57;  
 DT 01-MAR-2002 (TrEMBLrel. 20, Created)  
 DT 01-MAR-2002 (TrEMBLrel. 20, Last sequence update)  
 DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)  
 DE Gyra (Fragment).  
 GN GYRA.  
 OS Proteus mirabilis.  
 OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;  
 OC Enterobacteriaceae; Proteus.  
 OX NCBI\_TaxID=584;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=ATCC 29306;  
 RA Weigel L.M., Anderson G.J., Tenover F.C.;  
 RT "Gyrase and parC Mutations Associated with Fluoroquinolone Resistance in  
 RT Proteus mirabilis";  
 RL Submitted (JUL-2001) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; AF397169; AAL57340.1; -;  
 DR InterPro; IPR002205; DNA\_topoisom.  
 DR Pfam; PF00521; DNA\_topoisom; 1.  
 DR ProDom; PD000742; DNA\_topoisom; 1.  
 DR SMART; SM00434; TOP4C; 1.  
 FT NON\_TER 1  
 FT NON\_TER 98  
 SQ SEQUENCE 98 AA; 10963 MW; 052CD25D0FA5EF1F CRC64;

Query Match 66.7%; Score 28; DB 2; Length 98;  
 Best Local Similarity 62.5%; Pred. No. 2.5e+02;  
 Matches 5; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 2 VLMAMNII 9  
 |||:|

Db 8 VLFAMNVL 15  
 |||:|

Tue Feb 17 11:55:48 2004

us-09-900-147-2.rspt

Query Match 66.7%; Score 28; DB 8; Length 98;  
 Best Local Similarity 62.5%; Pred. No. 2.5e+02;  
 Matches 5; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

OY 1 NVLMAMNI 8  
 Db 44 NTLMALNM 51

## RESULT 12

ID O21712 PRELIMINARY; PRT; 98 AA.  
 AC O21712;  
 DT 01-JAN-1998 (TREMELrel. 05, Created)  
 DT 01-JAN-1998 (TREMELrel. 05, Last sequence update)  
 DT 01-JUN-2002 (TREMELrel. 21, Last annotation update)  
 DE NADH dehydrogenase subunit 4L.  
 GN ND4L.  
 OS Pygathrix roxellana (golden snub-nosed monkey).  
 OG Mitochondrion.  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Euthera; Primates; Catarrhini; Cercopitheidae; Colobinae;  
 OC Pygathrix.  
 OC NCBI\_TaxID=61622;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=CUJ;  
 RA Wang W., Forstner M.R.J., Zhang Y., Lui Z., Wei Y., Huang H., Hu H.,  
 RA Xie Y., Wu D., Melnick D.J.;  
 RA "A Phylogeny of Chinese Leaf Monkeys Using Mitochondrial ND3-ND4 Gene  
 RT Sequences";  
 RL Int. J. Primatol. 0:0-0(1997).  
 DR EMBL; U92962; AAD04663.1; -;  
 DR InterPro; IPR001133; Oxidored4L.  
 DR InterPro; IPR003214; Oxidored4L.  
 DR Pfam; PF00420; Oxidored\_q2; 1.  
 DR ProDom; PD000359; Oxidred4L; 1.  
 KW Mitochondrion.  
 SQ SEQUENCE 98 AA; 10856 MW; 2B2F172767A7DCA9 CRC64;

Query Match 66.7%; Score 28; DB 8; Length 98;  
 Best Local Similarity 62.5%; Pred. No. 2.5e+02;  
 Matches 5; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

OY 1 NVLMAMNI 8  
 Db 44 NTLMALNM 51

## RESULT 13

ID O21846 PRELIMINARY; PRT; 98 AA.  
 AC O21846;  
 DT 01-JAN-1998 (TREMELrel. 05, Created)  
 DT 01-JAN-1998 (TREMELrel. 05, Last sequence update)  
 DT 01-JUN-2002 (TREMELrel. 21, Last annotation update)  
 DE NADH dehydrogenase subunit 4L.  
 GN ND4L.  
 OS Pygathrix roxellana (golden snub-nosed monkey).  
 OG Mitochondrion.  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Euthera; Primates; Catarrhini; Cercopitheidae; Colobinae;  
 OC Pygathrix.  
 OC NCBI\_TaxID=61622;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=CUJ3, and CUJ1;  
 RA Wang W., Forstner M.R.J., Zhang Y., Lui Z., Wei Y., Huang H., Hu H.,  
 RA Xie Y., Wu D., Melnick D.J.;  
 RA Int. J. Primatol. 0:0-0(1997).  
 RL EMBL; U92963; AAD04666.1; -;  
 DR EMBL; U92961; AAD04660.1; -;

RESULT 10  
 O21757 PRELIMINARY; PRT; 98 AA.  
 AC O21757;  
 DT 01-JAN-1998 (TREMELrel. 05, Created)  
 DT 01-JAN-1998 (TREMELrel. 05, Last sequence update)  
 DT 01-JUN-2002 (TREMELrel. 21, Last annotation update)  
 DE NADH dehydrogenase subunit 4L.  
 GN ND4L.  
 OS Pygathrix bieti (Black snub-nosed monkey) (Rhinopithecus bieti).  
 OG Mitochondrion.  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Euthera; Primates; Catarrhini; Cercopitheidae; Colobinae;  
 OC Pygathrix.  
 OC NCBI\_TaxID=61621;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=DJ7 AND DJ1 AND DJ4, and DJ5;  
 RA Wang W., Forstner M.R.J., Zhang Y., Lui Z., Wei Y., Huang H., Hu H.,  
 RA Xie Y., Wu D., Melnick D.J.;  
 RA Int. J. Primatol. 0:0-0(1997).  
 DR EMBL; U92960; AAD08834.1; -;  
 DR EMBL; U92951; AAD04678.1; -;  
 DR EMBL; U92957; AAD08825.1; -;  
 DR EMBL; U92958; AAD08828.1; -;  
 DR InterPro; IPR001133; Oxidored4L.  
 DR InterPro; IPR003214; Oxidred4L.  
 DR Pfam; PF00420; Oxidored\_q2; 1.  
 DR ProDom; PD000359; Oxidred4L; 1.  
 KW Mitochondrion.  
 SQ SEQUENCE 98 AA; 10894 MW; CC796F277268368A CRC64;

Query Match 66.7%; Score 28; DB 8; Length 98;  
 Best Local Similarity 62.5%; Pred. No. 2.5e+02;  
 Matches 5; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

OY 1 NVLMAMNI 8  
 Db 44 NTLMALNM 51

## RESULT 11

ID O21705 PRELIMINARY; PRT; 98 AA.  
 AC O21705;  
 DT 01-JAN-1998 (TREMELrel. 05, Created)  
 DT 01-JAN-1998 (TREMELrel. 05, Last sequence update)  
 DT 01-JUN-2002 (TREMELrel. 21, Last annotation update)  
 DE NADH dehydrogenase subunit 4L.  
 GN ND4L.  
 OS Pygathrix bieti (Black snub-nosed monkey) (Rhinopithecus bieti).  
 OG Mitochondrion.  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Euthera; Primates; Catarrhini; Cercopitheidae; Colobinae;  
 OC Pygathrix.  
 OC NCBI\_TaxID=61621;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=DJ3;  
 RA Wang W., Forstner M.R.J., Zhang Y., Lui Z., Wei Y., Huang H., Hu H.,  
 RA Xie Y., Wu D., Melnick D.J.;  
 RA "A Phylogeny of Chinese Leaf Monkeys Using Mitochondrial ND3-ND4 Gene  
 RT Sequences";  
 RL Int. J. Primatol. 0:0-0(1997).  
 DR EMBL; U92956; AAD08822.1; -;  
 DR InterPro; IPR001133; Oxidored4L.  
 DR InterPro; IPR003214; Oxidred4L.  
 DR Pfam; PF00420; Oxidored\_q2; 1.  
 DR ProDom; PD000359; Oxidred4L; 1.  
 KW Mitochondrion.  
 SQ SEQUENCE 98 AA; 10902 MW; 78796F27726E9AE0 CRC64;

```

DR InterPro; IPR001133; Oxidored4L.
DR InterPro; IPR003214; Oxidred4L.
DR Pfam; PF00420; Oxidored_G2; 1.
DR ProDom; PD000359; Oxidred4L; 1.
KW Mitochondrion.
SQ SEQUENCE 98 AA; 10923 MW; 71683E3FF63B8572 CRC64;

Query Match      66.7%; Score 28; DB 8; Length 98;
Best Local Similarity 62.5%; Pred. No. 2.5e+02;
Matches 5; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Qy 1 NVLMANNI 8
Db 44 NTLMLNM 51

RESULT 14
ID Q8CPD5 PRELIMINARY; PRT; 47 AA.
AC Q8CPD5;
DT 01-MAR-2003 (TrEMBLrel. 23, Created)
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE Truncated transposase.
GN SEI005.
OS Staphylococcus epidermidis.
OC Bacteria; Firmicutes; Bacillales; Staphylococcus.
OX NCBI_TaxID=1282;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC 12228;
RA Zhang Y., Ren S., Li H., Fu G., Lu L., Lu G., Jia J., Tu Y., Qin Z.,
RA Chen Z., Wen Y.;
RL Submitted (NOV-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AEO16747; AAC04602.1; -.
KW Complete proteome.
SQ SEQUENCE 47 AA; 5498 MW; 170FAA03C09426DD CRC64;

Query Match      64.3%; Score 27; DB 16; Length 47;
Best Local Similarity 85.7%; Pred. No. 2.1e+02;
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 2 VLMANNI 8
Db 32 VLMANNI 38

RESULT 15
Q8X8C1
ID Q8X8C1 PRELIMINARY; PRT; 81 AA.
AC Q8X8C1;
DT 01-MAR-2002 (TrEMBLrel. 20, Created)
DT 01-MAR-2002 (TrEMBLrel. 20, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE Orf, hypothetical protein.
GN YIIIE OR Z5431 OR ECS4815.
OS Escherichia coli O157:H7.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Escherichia.
OX NCBI_TaxID=83334;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=O157:H7 / EDL933 / ATCC 700927;
RX MEDLINE=21074935; PubMed=11206551;
RA Perna N.T., Plunkett G. III, Burland V., Mau B., Glasner J.D.,
RA Rose D.J., Mayhew G.F., Evans P.S., Gregor J., Kirkpatrick H.A.,
RA Postel G., Hackett J., Klink S., Boutin A., Shao Y., Miller L.,
RA Grobeck E.J., Davis N.W., Lim A., Dimalanta E.T., Potamousis K.,
RA Apodaca J., Anantharaman T.S., Lin J., Yen G., Schwartz D.C.,
RA Welch R.A., Blattner F.R.;
RT "Genome sequence of enterohaemorrhagic Escherichia coli O157:H7.";
RL Nature 409:529-533 (2001).
RN [2]

RP SEQUENCE FROM N.A.
RC STRAIN=O157:H7 / RIMD 0509952;
RX MEDLINE=21156231; PubMed=11258796;
RA Havaashi T., Makino K., Ohnishi M., Kurokawa K., Ishii K., Yokoyama K.,
RA Han C.-G., Ohtsubo E., Nakayama K., Murata T., Tanaka M., Tobe T.,
RA Iida T., Takami H., Honda T., Sasakawa C., Ogasawara N., Yasunaga T.,
RA Kuhara S., Shiba T., Hattori M., Shinagawa H.;
RT "Complete genome sequence of enterohaemorrhagic Escherichia coli O157:H7 and genomic comparison with a laboratory strain K-12.";
RL DNA Res. 8:11-22 (2001).
DR EMBL; AE005620; BAG59082.1; -.
DR EMBL; AP002567; BAB38238.1; -.
DR InterPro; IPR002145; HTH_CopG.
DR Pfam; PF01402; HTH_4; 1.
KW Complete proteome.
SQ SEQUENCE 81 AA; 9382 MW; E45072094F7A729A CRC64;

Query Match      64.3%; Score 27; DB 16; Length 81;
Best Local Similarity 62.5%; Pred. No. 3.4e+02;
Matches 5; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Qy 2 VLMANNI 9
Db 8 VLMANNI 15

RESULT 16
Q8PBE9 PRELIMINARY; PRT; 81 AA.
ID Q8PBE9;
AC Q8PBE9;
DT 01-MAR-2003 (TrEMBLrel. 23, Created)
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE Hypothetical protein yIIIE.
GN YIIIE OR C4840.
OS Escherichia coli O6.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Escherichia.
OX NCBI_TaxID=217992;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=O6:H1 / CFT073 / ATCC 700928;
RX MEDLINE=22388234; PubMed=12471157;
RA Welch R.A., Burland V., Plunkett G. III, Redford P., Roesch P.,
RA Rasko D., Buckles E.L., Liou S.-R., Boutin A., Hackett J., Stroud D.,
RA Mayhew G.F., Rose D.J., Zhou S., Schwartz D.C., Perna N.T.,
RA Mobley H.L.T., Donnenberg M.S., Blattner F.R.;
RT "Extensive mosaic structure revealed by the complete genome sequence of uropathogenic Escherichia coli.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:17020-17024 (2002).
DR EMBL; AE016770; AAN83269.1; -.
KW Hypothetical protein; Complete proteome.
SQ SEQUENCE 81 AA; 9452 MW; 37E572149DF160BB CRC64;

Query Match      64.3%; Score 27; DB 16; Length 81;
Best Local Similarity 62.5%; Pred. No. 3.4e+02;
Matches 5; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Qy 2 VLMANNI 9
Db 8 VLMANNI 15

RESULT 17
Q8AWZ8 PRELIMINARY; PRT; 99 AA.
ID Q8AWZ8;
AC Q8AWZ8;
DT 01-MAR-2003 (TrEMBLrel. 23, Created)
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE Voltage-gated potassium channel subunit MiRP2.
GN KCNE3.
OS Xenopus laevis (African clawed frog).

```

Tue Feb 17 11:55:48 2004

us-09-900-147-2.rspt

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipidae;  
 OC Xenopodinae; Xenopus.  
 OX NCBI\_TaxID=8355;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Anantharam A., Lewis A., Panaghie G., McCrossan Z.A., Lerner D.J.,  
 RA Abbott G.W.; Xenopus oocyte MiRps govern mammalian K+ channel function  
 RT "Endogenous Xenopus oocyte variability."  
 RL Submitted (SEP-2002) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; AF545501; AAN77245.1; -  
 KW Ionic channel.  
 SQ SEQUENCE 99 AA; 11332 MW; 53A06B045F73FAE3 CRC64;

Query Match 64.3%; Score 27; DB 13; Length 99;  
 Best Local Similarity 71.4%; Pred. No. 4.1e+02;  
 Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 NVLMANN 7  
 DB 16 NVLLAIN 22

RESULT 18  
 Q97YM7 PRELIMINARY; PRT; 99 AA.

AC Q97YM7;  
 DT 01-OCT-2001 (TrEMBLrel. 18, Created)  
 DT 01-OCT-2001 (TrEMBLrel. 18, Last sequence update)  
 DT 01-OCT-2001 (TrEMBLrel. 18, Last annotation update)  
 DE Hypothetical protein SSO1296.  
 GN SSO1296.  
 OS Sulfolobus solfataricus.  
 OC Archaea; Crenarchaeota; Thermoprotei; Sulfolobales; Sulfolobaceae;  
 OC Sulfolobus.  
 OX NCBI\_TaxID=2287;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=ATCC 35092 / DSM 1617 / P2;  
 RX MEDLINE=21332296; PubMed=11427726;  
 RA She Q., Singh R.K., Confalonieri F., Zivanovic Y., Allard G.,  
 RA Awayez M.J., Chan-Weiher C.C.-Y., Clausen I.G., Curtis B.A.,  
 RA De Moors A., Ebraus G., Fletcher C., Gordon P.M.K.,  
 RA Heikamp-de Jong I., Jeffries A.C., Kozera C.J., Medina N., Peng X.,  
 RA Thi-Ngoc H.P., Redder P., Schenk M.E., Theriault C., Tolettrup N.,  
 RA Charlebois R.L., Doolittle W.F., Duguet M., Gaasterland T.,  
 RA Gartett R.A., Ragan M.A., Senses C.W., Van der Oost J.;  
 RT "The complete genome of the crenarchaeon Sulfolobus solfataricus P2."  
 RL Proc. Natl. Acad. Sci. U.S.A. 98:7835-7840(2001).  
 DR EMBL; AE006744; AAK41533.1; -  
 KW Hypothetical protein; Complete proteome.  
 SQ SEQUENCE 99 AA; 11427 MW; A9DD456CB7BB7F7 CRC64;

Query Match 64.3%; Score 27; DB 17; Length 99;  
 Best Local Similarity 44.4%; Pred. No. 4.1e+02;  
 Matches 4; Conservative 4; Mismatches 1; Indels 0; Gaps 0;

QY 1 NVLMANNII 9  
 DB 75 NIIISMNAI 83

RESULT 19  
 ID Q8CQ16 PRELIMINARY; PRT; 40 AA.

AC Q8CQ16;  
 DT 01-MAR-2003 (TrEMBLrel. 23, Created)  
 DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)  
 DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)  
 DE Truncated transposase.  
 GN SSO083.  
 OS Staphylococcus epidermidis.

OC Bacteria; Firmicutes; Bacillales; Staphylococcus.  
 OX NCBI\_TaxID=1282;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=ATCC 12228;  
 RA Zhang Y., Ren S., Li H., Fu G., Lu L., Lu G., Jia J., Tu Y., Qin Z.,  
 RA Chen Z., Wen Y.;  
 RL Submitted (NOV-2002) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; AE016744; AA003660.1; -  
 KW Complete proteome.  
 SQ SEQUENCE 40 AA; 4792 MW; FEEL3BA49163D7F9 CRC64;

Query Match 61.9%; Score 26; DB 16; Length 40;  
 Best Local Similarity 71.4%; Pred. No. 3e+02;  
 Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 2 VLMANNI 8  
 DB 1 MLMALNI 7

RESULT 20  
 Q9GCA4 PRELIMINARY; PRT; 50 AA.

AC Q9GCA4;  
 DT 01-MAR-2001 (TrEMBLrel. 16, Created)  
 DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)  
 DT 01-JUN-2001 (TrEMBLrel. 17, Last annotation update)  
 DE NADH dehydrogenase subunit 3 (fragment).  
 GN ND3.  
 OS Samia cynthia ricini (Indian eri silkworm).  
 OC Mitochondrion.  
 OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;  
 OC Neoptera; Endopterygota; Lepidoptera; Glossata; Bombycoidea;  
 OC Saturniidae; Saturniinae; Attacini; Samia.  
 OX NCBI\_TaxID=30247;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Wei Z., Zhao Q., Zhang Z., Wang Z., He J.;  
 RT "Eri Silkworm Mitochondrial Cytochrome Oxidase Subunit III, trna-Gly  
 RT and Partial NADH dehydrogenase Subunit 3 genes."  
 RL Submitted (JUN-2000) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; AF288145; AAG01169.1; -  
 DR InterPro; IPR000440; Oxidored\_q4.  
 DR Pfam; PF00507; oxidored\_q4; 1.  
 KW Mitochondrion.  
 FT NON TER 50 50  
 SQ SEQUENCE 50 AA; 5711 MW; BB7A3D4768A49F2D CRC64;

Query Match 61.9%; Score 26; DB 8; Length 50;  
 Best Local Similarity 55.6%; Pred. No. 3.6e+02;  
 Matches 5; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 1 NVLMANNII 9  
 DB 15 NLMLILSII 23

RESULT 21  
 Q8X3F6 PRELIMINARY; PRT; 53 AA.

AC Q8X3F6;  
 DT 01-MAR-2002 (TrEMBLrel. 20, Created)  
 DT 01-MAR-2002 (TrEMBLrel. 20, Last sequence update)  
 DT 01-MAR-2002 (TrEMBLrel. 20, Last annotation update)  
 DE Hypothetical protein ECs3484.  
 GN ECs3484.  
 OS Escherichia coli O157:H7.  
 OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;  
 OC Enterobacteriaceae; Escherichia.  
 OX NCBI\_TaxID=83334;  
 RN [1]  
 RP SEQUENCE FROM N.A.

```

RC STRAIN=0157:H7 / RIMD 0509952;
RX MEDLINE=21156231; PubMed=11258796;
RA Hayaishi T., Makino K., Ohnishi M., Kurokawa K., Ishii K., Yokoyama K.,
RA Han C.-G., Ohtsubo E., Nakayama K., Murata T., Tanaka M., Tobe T.,
RA Lida T., Takami H., Honda T., Sasakawa C., Ogasawara N., Yasunaga T.,
RA Kuhara S., Shiba T., Hattori M., Shinagawa H.;
RT "Complete genome sequence of enterohemorrhagic Escherichia coli
RT O157:H7 and genomic comparison with a laboratory strain K-12."
RL DNA Res. 8:11-22(2001).
DR EMBL; AF002562; BAB36907.1; -.
KW Hypothetical protein.
SQ SEQUENCE 53 AA; 6018 MW; 653FB85407E4B98F CRC64;

Query Match 61.9%; Score 26; DB 16; Length 53;
Best Local Similarity 50.0%; Pred. No. 3.8e+02;
Matches 4; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 1 NVLMANNII 8
Db |::|::|
43 NIVAAWNL 50

RESULT 22
Q82697 PRELIMINARY; PRT; 73 AA.
AC Q82697;
DT 01-MAR-2002 (TrEMBLrel. 20, Created)
DT 01-MAR-2002 (TrEMBLrel. 20, Last sequence update)
DT 01-MAR-2002 (TrEMBLrel. 20, Last annotation update)
DE Putative membrane protein.
GN STY1908.
OS Salmonella typhi.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Salmonella.
OX NCBI_TaxID=601;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CT18;
RX MEDLINE=21534947; PubMed=11677608;
RA Parkhill J., Dougan G., James K.D., Thomson N.R., Pickard D., Wain J.,
RA Churcher C., Mungall K.L., Bentley S.D., Holden M.T.G., Sebahia M.,
RA Baker S., Basham D., Brooks K., Chillingworth T., Connor P.,
RA Cronin A., Davis P., Davies R.M., Dowd L., White N., Farrar J.,
RA Felwell T., Hamlin N., Haque A., Hien T.T., Holroyd S., Jagels K.,
RA Krogh A., Larsen T.S., Leather S., Moule S., O'Gaora P., Parry C.,
RA Quail M., Rutherford K., Simmonds M., Skelton J., Stevens K.,
RA Whitehead S., Barrell B.G.;
RT "Complete genome sequence of a multiple drug resistant Salmonella
RT enterica serovar Typhi CT18."
RL Nature 413:848-852(2001).
DR EMBL; AL627271; CAD02137.1; -.
KW Hypothetical protein; Complete proteome.
SQ SEQUENCE 73 AA; 7846 MW; 74C42318C53CCA58 CRC64;

Query Match 61.9%; Score 26; DB 16; Length 73;
Best Local Similarity 44.4%; Pred. No. 5.1e+02;
Matches 4; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 1 NVLMANNII 9
Db |::|::|
11 NVLMGLGLV 19

RESULT 23
Q8W9T7 PRELIMINARY; PRT; 84 AA.
AC Q8W9T7;
DT 01-MAR-2002 (TrEMBLrel. 20, Created)
DT 01-MAR-2002 (TrEMBLrel. 20, Last sequence update)
DT 01-OCT-2002 (TrEMBLrel. 22, Last annotation update)
DE NADH dehydrogenase subunit 4L (Fragment).
GN ND4L.
OS Leptasterias polaris.

```

```

OG Mitochondrion.
OC Eukaryota; Metazoa; Echinodermata; Eleutherozoa; Asterozoa;
OC Asteroidea; Forcipulatacea; Forcipulata; Asteriidae; Leptasterias.
OX NCBI_TaxID=59564;
RN [1]
RP SEQUENCE FROM N.A.
RA Hrinkevich A.W., Rocha-Olivares A., Foltz D.W.;
RT "Phylogenetic analysis of molecular lineages in a species-rich
RT subgenus of sea stars (Leptasterias subgenus Hexasterias).";
RL Am. Zoologist 40:365-374(2000).
RN [2]
RP SEQUENCE FROM N.A.
RA Hrinkevich A.W., Rocha-Olivares A., Foltz D.W.;
RL Submitted (MAY-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF162110; AAL61885.1; -.
DR InterPro; IPR001133; Oxidored_4L.
DR Pfam; PF00420; Oxidored_q2; 1.
KW Mitochondrion.
FT NON TER 84 84
SQ SEQUENCE 84 AA; 9151 MW; AADAFD859FE82405 CRC64;

Query Match 61.9%; Score 26; DB 8; Length 84;
Best Local Similarity 44.4%; Pred. No. 5.7e+02;
Matches 4; Conservative 4; Mismatches 1; Indels 0; Gaps 0;

QY 1 NVLMANNII 9
Db |::|::|
54 NILLNNLI 62

RESULT 24
Q82P30 PRELIMINARY; PRT; 92 AA.
ID Q82P30;
AC Q82P30;
DT 01-MAR-2002 (TrEMBLrel. 20, Created)
DT 01-MAR-2002 (TrEMBLrel. 20, Last sequence update)
DT 01-MAR-2002 (TrEMBLrel. 20, Last annotation update)
DE Putative inner membrane protein.
GN YCHH OR STMI782.
OS Salmonella typhimurium.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Salmonella.
OX NCBI_TaxID=602;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=LT2 / SGSC1412 / ATCC 700720;
RX MEDLINE=21534948; PubMed=11677609;
RA McClelland M., Sanderson K.E., Spieth J., Clifton S.W., Latreille P.,
RA Courtney L., Porwollik S., Ali J., Dante M., Du P., Hou S., Layman D.,
RA Leonard S., Nguyen C., Scott K., Holmes A., Grewal N., Mulvaney E.,
RA Ryan E., Sun H., Florea L., Miller W., Stoneking T., Nhan M.,
RA Waterston R., Wilson R.K.;
RT "Complete genome sequence of Salmonella enterica serovar Typhimurium
RT LT2."
RL Nature 413:852-856(2001).
DR EMBL; AE008779; AAL20697.1; -.
KW Hypothetical protein; Complete proteome.
SQ SEQUENCE 92 AA; 10469 MW; 270E355575AFF3F4 CRC64;

Query Match 61.9%; Score 26; DB 16; Length 92;
Best Local Similarity 44.4%; Pred. No. 6.2e+02;
Matches 4; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 1 NVLMANNII 9
Db |::|::|
11 NVLMGLGLV 19

RESULT 25
Q96DE8 PRELIMINARY; PRT; 95 AA.
ID Q96DE8;
AC Q96DE8;
DT 01-DEC-2001 (TrEMBLrel. 19, Created)

```

DT	01-DEC-2001 (TrEMBLrel. 19, Last sequence update)	AC	Q8YG45; 2002 (TrEMBLrel. 20, Created)
DT	01-OCT-2002 (TrEMBLrel. 22, Last annotation update)	DT	01-MAR-2002 (TrEMBLrel. 20, Last sequence update)
DE	Similar to 26S proteasome-associated padi homolog.	DT	01-MAR-2002 (TrEMBLrel. 20, Last annotation update)
OS	Homo sapiens (Human)	DN	Hypothetical protein BMEI1317.
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;	GN	BMEI1317.
OC	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.	OS	Brucella melitensis.
OC	NCBI_TaxID=9606;	OC	Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;
RN	[1]	OC	Brucellaceae; Brucella.
RP	SEQUENCE FROM N.A.	OX	NCBI_TaxID=29459;
RC	TISSUE=Lung;	RN	[1]
RA	Strausberg R.;	RP	SEQUENCE FROM N.A.
RL	Submitted (JUN-2001) to the EMBL/GenBank/DBJ databases.	RC	STRAIN=16M / ATCC 23456 / Biotype 1;
DR	EMBL; BC009524; AAH09524.1; -	RX	MEDLINE=20020109; PubMed=11756688;
DR	InterPro; IPR003639; Mov34_1.	RA	DelVecchio V.G., Kapatral V., Redkar R.J., Patra G., Mijer C., Los T.,
DR	ProDom; PD363422; Mov34_1_1.	RA	Ivanova N., Anderson I., Bhattacharya A., Lykidis A., Reznik G.,
KW	Proteasome.	RA	Jablonski L., Larsen N., D'Souza M., Bernal A., Mazur M., Goltsman E.,
SQ	SEQUENCE 95 AA; 11009 MW; D70270AD00529A18 CRC64;	RA	Selkov E., Elzer P.H., Hagius S., O'Callaghan D., Letesson J.-J.,
		RA	Hasekorn R., Kyripides N., Overbeek R.;
Query Match	61.9%; Score 26; DB 4; Length 95;	RT	"The genome sequence of the facultative intracellular pathogen
Best Local Similarity	55.6%; Pred. No. 6.4e+02;	RT	Brucella melitensis.";
Matches	5; Conservative 2; Mismatches 2; Indels 0; Gaps 0;	RL	Proc. Natl. Acad. Sci. U.S.A. 99:443-448(2002).
Qy	1 NLMANNII 9	DR	EMBL; AE009569; AAL52498.1; -
Db	74 DVLMTSNIV 82	KW	Hypothetical protein; Complete proteome.
		SQ	SEQUENCE 98 AA; 10766 MW; 99892EBIDEAC40D CRC64;
RESULT 26			
Q8LWU8	PRELIMINARY; PRT; 98 AA.	Query Match	61.9%; Score 26; DB 16; Length 98;
AC	Q8LWU8;	Best Local Similarity	55.6%; Pred. No. 6.6e+02;
DT	01-OCT-2002 (TrEMBLrel. 22, Created)	Matches	5; Conservative 3; Mismatches 1; Indels 0; Gaps 0;
DT	01-OCT-2002 (TrEMBLrel. 22, Last sequence update)	Qy	1 NLMANNII 9
DT	01-MAR-2003 (TrEMBLrel. 23, Last annotation update)	Db	25 NRLVAINIV 33
DE	NADH dehydrogenase subunit 4L.		
GN	NADH4L.	RESULT 28	
OS	Macroscelides proboscideus (Short-eared elephant shrew).	Q8G1S9	PRELIMINARY; PRT; 98 AA.
OG	Mitochondrion.	AC	Q8G1S9;
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;	DT	01-MAR-2003 (TrEMBLrel. 23, Created)
OC	Mammalia; Eutheria; Macroscelidae; Macroscelididae; Macroscelides.	DT	01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
OC	NCBI_TaxID=29082;	DT	01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
RN	[1]	DE	Hypothetical protein.
RP	SEQUENCE FROM N.A.	GN	BR0625.
RA	Aranson U.;	OS	Brucella suis.
RL	Submitted (DEC-2001) to the EMBL/GenBank/DBJ databases.	OC	Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;
RP	[2]	OC	Brucellaceae; Brucella.
RP	SEQUENCE FROM N.A.	OX	NCBI_TaxID=29461;
RX	MEDLINE=22056029; PubMed=12034869;	RN	[1]
RA	Aranson U., Adegoke J.A., Bodin K., Born E.W., Esa Y.B., Gullberg A.,	RP	SEQUENCE FROM N.A.
RA	Nilsson M., Short R.V., Xu X., Janke A.;	RC	STRAIN=1330 / Biovar 1;
RT	"Mammalian mitogenomic relationships and the root of the eutherian	RX	MEDLINE=22247741; PubMed=12271122;
RT	tree.";	RA	Paulsen I.T., Seshadri R., Nelson K.E., Eisen J.A., Heidelberg J.F.,
RL	Proc. Natl. Acad. Sci. U.S.A. 99:8151-8156(2002).	RA	Read T.D., Dodson R.J., Umayam L., Brinkac L.M., Beanan M.J.,
DR	EMBL; AJ421452; CAD13403.1; -	RA	Daugherty S.C., Debay R.T., Durkin A.S., Kolonay J.F., Madupu R.,
DR	InterPro; IPR001133; Oxidored4L.	RA	Nelson W.C., Ayodeji B., Kraul M., Shetty J., Malek J., Van Aken S.E.,
DR	InterPro; IPR003214; Oxidored4L.	RA	Riedmuller S., Tettelin H., Gill S.R., White O., Salzberg S.L.,
DR	Pfam; PF00420; Oxidored4_2; 1.	RA	Hoover D.L., Lindler L.E., Halling S.M., Boyle S.M., Fraser C.M.;
DR	ProDom; PD000359; Oxidored4L; 1.	RT	"The Brucella suis genome reveals fundamental similarities between
KW	Mitochondrion.	RT	animal and plant pathogens and symbionts.";
SQ	SEQUENCE 98 AA; 10952 MW; 2B979C6CED115DFC CRC64;	RL	Proc. Natl. Acad. Sci. U.S.A. 99:13148-13153(2002).
		DR	EMBL; AE014370; AAN29554.1; -
Query Match	61.9%; Score 26; DB 8; Length 98;	DR	TIGR; BR0625; -
Best Local Similarity	71.4%; Pred. No. 6.6e+02;	KW	Hypothetical protein; Complete proteome.
Matches	5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;	SQ	SEQUENCE 98 AA; 10766 MW; 99892EBIDEAC40D CRC64;
Qy	3 LMANII 9	Query Match	61.9%; Score 26; DB 16; Length 98;
Db	3 LMTMNL 9	Best Local Similarity	55.6%; Pred. No. 6.6e+02;
		Matches	5; Conservative 3; Mismatches 1; Indels 0; Gaps 0;
RESULT 27		Qy	1 NLMANNII 9
Q8YG45	PRELIMINARY; PRT; 98 AA.	Db	25 NRLVAINIV 33
ID	Q8YG45		



```

RESULT 29
Q9XNR1
ID Q9XNR1 PRELIMINARY; PRT; 99 AA.
AC Q9XNR1;
DT 01-NOV-1999 (TREMBLrel. 12, Created)
DT 01-NOV-1999 (TREMBLrel. 12, Last sequence update)
DT 01-OCT-2002 (TREMBLrel. 22, Last annotation update)
DE NADH dehydrogenase subunit 4 (Fragment).
OS Boophilus microplus (Cattle tick).
OG Mitochondrion.
OC Eukaryota; Metazoa; Arthropoda; Chelicerata; Arachnida; Acari;
OC Parasitiformes; Ixodida; Ixodidae; Boophilus.
OX NCBI_TaxID=6941;
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN=N.
RX MEDLINE=99297341; PubMed=10368952;
RA Campbell N.J.H., Barker S.C.;
RT "The novel mitochondrial gene arrangement of the cattle tick,
RT Boophilus microplus: fivefold tandem repetition of a coding region.";
RL Mol. Biol. Evol. 16:732-740(1999).
DR EMBL, AF110610; AAD28378.1; -.
KW Mitochondrion.
FT NON TER 99
SQ SEQUENCE 99 AA; 11987 MW; 7AEAD5FB76C500 CRC64;

Query Match 61.9%; Score 26; DB 8; Length 99;
Best Local Similarity 50.0%; Pred. No. 6.6e+02;
Matches 4; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 2 VLMAMNII 9
Db 86 ILMTWLL 93

RESULT 30
Q9MIJ8
ID Q9MIJ8 PRELIMINARY; PRT; 21 AA.
AC Q9MIJ8;
DT 01-OCT-2000 (TREMBLrel. 15, Created)
DT 01-OCT-2000 (TREMBLrel. 15, Last sequence update)
DT 01-OCT-2000 (TREMBLrel. 15, Last annotation update)
DE NADH dehydrogenase subunit III (Fragment).
OS Cicindela wagneri.
OG Mitochondrion.
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Coleoptera; Adephaga; Cicindelidae;
OC Cicindela.
OX NCBI_TaxID=93367;
RN [1]
RP SEQUENCE FROM N.A.
RA Barracough T.G., Hogan J.E., Vogler A.P.;
RT "Testing whether ecological factors promote cladogenesis in a group of
RT tiger beetles (Coleoptera: Cicindelidae).";
RL Proc. R. Soc. Lond., B, Biol. Sci. 0:0-0(2000).
DR EMBL, AF133025; AAF89120.1; -.
KW Mitochondrion.
FT NON TER 1
FT NON TER 21
SQ SEQUENCE 21 AA; 2215 MW; EB3165AA2FAEBD45 CRC64;

Query Match 59.5%; Score 25; DB 8; Length 21;
Best Local Similarity 50.0%; Pred. No. 2.7e+02;
Matches 4; Conservative 4; Mismatches 0; Indels 0; Gaps 0;

QY 2 VLMAMNII 9
Db 5 IIALSII 12

RESULT 31
Q9MI15
ID Q9MI15 PRELIMINARY; PRT; 21 AA.
AC Q9MI15;
DT 01-OCT-2000 (TREMBLrel. 15, Created)
DT 01-OCT-2000 (TREMBLrel. 15, Last sequence update)

```

```

Q9MI19
ID Q9MI19 PRELIMINARY; PRT; 21 AA.
AC Q9MI19;
DT 01-OCT-2000 (TREMBLrel. 15, Created)
DT 01-OCT-2000 (TREMBLrel. 15, Last sequence update)
DT 01-OCT-2000 (TREMBLrel. 15, Last annotation update)
DE NADH dehydrogenase subunit III (Fragment).
OS Cicindela cuprascens.
OG Mitochondrion.
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Coleoptera; Adephaga; Cicindelidae;
OC Cicindela.
OX NCBI_TaxID=93357;
RN [1]
RP SEQUENCE FROM N.A.
RA Barracough T.G., Hogan J.E., Vogler A.P.;
RT "Testing whether ecological factors promote cladogenesis in a group of
RT tiger beetles (Coleoptera: Cicindelidae).";
RL Proc. R. Soc. Lond., B, Biol. Sci. 0:0-0(2000).
DR EMBL, AF133030; AAF89129.1; -.
KW Mitochondrion.
FT NON TER 21
SQ SEQUENCE 21 AA; 2247 MW; EB3165AA3EAFAD45 CRC64;

Query Match 59.5%; Score 25; DB 8; Length 21;
Best Local Similarity 50.0%; Pred. No. 2.7e+02;
Matches 4; Conservative 4; Mismatches 0; Indels 0; Gaps 0;

QY 2 VLMAMNII 9
Db 5 IIALSII 12

RESULT 32
Q9MIJ2
ID Q9MIJ2 PRELIMINARY; PRT; 21 AA.
AC Q9MIJ2;
DT 01-OCT-2000 (TREMBLrel. 15, Created)
DT 01-OCT-2000 (TREMBLrel. 15, Last sequence update)
DT 01-JUN-2001 (TREMBLrel. 17, Last annotation update)
DE NADH dehydrogenase subunit III (Fragment).
OS Cicindela sperata.
OG Mitochondrion.
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Coleoptera; Adephaga; Cicindelidae;
OC Cicindela.
OX NCBI_TaxID=93366;
RN [1]
RP SEQUENCE FROM N.A.
RA Barracough T.G., Hogan J.E., Vogler A.P.;
RT "Testing whether ecological factors promote cladogenesis in a group of
RT tiger beetles (Coleoptera: Cicindelidae).";
RL Proc. R. Soc. Lond., B, Biol. Sci. 0:0-0(2000).
DR EMBL, AF133028; AAF89126.1; -.
KW Mitochondrion.
FT NON TER 21
SQ SEQUENCE 21 AA; 2244 MW; EB2CA5AA3EAFAD45 CRC64;

Query Match 59.5%; Score 25; DB 8; Length 21;
Best Local Similarity 50.0%; Pred. No. 2.7e+02;
Matches 4; Conservative 4; Mismatches 0; Indels 0; Gaps 0;

QY 2 VLMAMNII 9
Db 5 IIALSII 12

RESULT 33
Q9MI15
ID Q9MI15 PRELIMINARY; PRT; 21 AA.
AC Q9MI15;
DT 01-OCT-2000 (TREMBLrel. 15, Created)
DT 01-OCT-2000 (TREMBLrel. 15, Last sequence update)

```

Tue Feb 17 11:55:48 2004

us-09-900-147-2.rspt

```

DT 01-OCT-2000 (TRENBLrel. 15, Last annotation update)
DE NADH dehydrogenase subunit III (Fragment).
OS Cicindela marutha.
OC Mitochondrion.
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Coleoptera; Adephaga; Cicindelidae;
OC Cicindela.
OX NCBI_TaxID=93364;
RN [1]
RN SEQUENCE FROM N.A.
RP Barracough T.G., Hogan J.E., Vogler A.P.;
RT "Testing whether ecological factors promote cladogenesis in a group of
RT tiger beetles (Coleoptera: Cicindelidae).";
RL Proc. R. Soc. Lond., B, Biol. Sci. 0:0-0(2000).
DR EMBL; AF133026; AAF89122.1; -.
KW Mitochondrion.
FT NON_TER 21
SQ SEQUENCE 21 AA; 2226 MW; EB2A641A3EAFAD45 CRC64;

Query Match 59.5%; Score 25; DB 8; Length 21;
Best Local Similarity 50.0%; Pred. No. 2.7e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 VLMAMNII 9
Db :||:|
5 IIMALSII 12

RESULT 34
Q9MIJ6 PRELIMINARY; PRT; 21 AA.
AC Q9MIJ6;
DT 01-OCT-2000 (TRENBLrel. 15, Created)
DT 01-OCT-2000 (TRENBLrel. 15, Last sequence update)
DT 01-OCT-2000 (TRENBLrel. 15, Last annotation update)
DE NADH dehydrogenase subunit III (Fragment).
OS Cicindela blanda.
OG Mitochondrion.
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Coleoptera; Adephaga; Cicindelidae;
OC Cicindela.
OX NCBI_TaxID=93356;
RN [1]
RN SEQUENCE FROM N.A.
RP Barracough T.G., Hogan J.E., Vogler A.P.;
RT "Testing whether ecological factors promote cladogenesis in a group of
RT tiger beetles (Coleoptera: Cicindelidae).";
RL Proc. R. Soc. Lond., B, Biol. Sci. 0:0-0(2000).
DR EMBL; AF133026; AAF89122.1; -.
KW Mitochondrion.
FT NON_TER 21
SQ SEQUENCE 21 AA; 2217 MW; EB23D25A3EAFAD45 CRC64;

Query Match 59.5%; Score 25; DB 8; Length 21;
Best Local Similarity 50.0%; Pred. No. 2.7e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 VLMAMNII 9
Db :||:|
5 IIMALSII 12

RESULT 35
Q9MIJ4 PRELIMINARY; PRT; 21 AA.
AC Q9MIJ4;
DT 01-OCT-2000 (TRENBLrel. 15, Created)
DT 01-OCT-2000 (TRENBLrel. 15, Last sequence update)
DT 01-OCT-2000 (TRENBLrel. 15, Last annotation update)
DE NADH dehydrogenase subunit III (Fragment).
OS Cicindela hirtillaris.
OG Mitochondrion.
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;

```

---

```

OC Neoptera; Endopterygota; Coleoptera; Adephaga; Cicindelidae;
OC Cicindela.
OX NCBI_TaxID=93360;
RN [1]
RN SEQUENCE FROM N.A.
RP Barracough T.G., Hogan J.E., Vogler A.P.;
RT "Testing whether ecological factors promote cladogenesis in a group of
RT tiger beetles (Coleoptera: Cicindelidae).";
RL Proc. R. Soc. Lond., B, Biol. Sci. 0:0-0(2000).
DR EMBL; AF133027; AAF89124.1; -.
KW Mitochondrion.
FT NON_TER 21
SQ SEQUENCE 21 AA; 2261 MW; EB3165AA3EBD3D45 CRC64;

Query Match 59.5%; Score 25; DB 8; Length 21;
Best Local Similarity 50.0%; Pred. No. 2.7e+02;
Matches 4; Conservative 4; Mismatches 0; Indels 0; Gaps 0;

QY 2 VLMAMNII 9
Db :||:|
5 IIMALSII 12

RESULT 36
Q9MIKO PRELIMINARY; PRT; 21 AA.
AC Q9MIKO;
DT 01-OCT-2000 (TRENBLrel. 15, Created)
DT 01-OCT-2000 (TRENBLrel. 15, Last sequence update)
DT 01-OCT-2000 (TRENBLrel. 15, Last annotation update)
DE NADH dehydrogenase subunit III (Fragment).
OS Cicindela gratiosa.
OG Mitochondrion.
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Coleoptera; Adephaga; Cicindelidae;
OC Cicindela.
OX NCBI_TaxID=93358;
RN [1]
RN SEQUENCE FROM N.A.
RP Barracough T.G., Hogan J.E., Vogler A.P.;
RT "Testing whether ecological factors promote cladogenesis in a group of
RT tiger beetles (Coleoptera: Cicindelidae).";
RL Proc. R. Soc. Lond., B, Biol. Sci. 0:0-0(2000).
DR EMBL; AF133024; AAF89118.1; -.
KW Mitochondrion.
FT NON_TER 21
SQ SEQUENCE 21 AA; 2243 MW; EB3165AA3F664D45 CRC64;

Query Match 59.5%; Score 25; DB 8; Length 21;
Best Local Similarity 50.0%; Pred. No. 2.7e+02;
Matches 4; Conservative 4; Mismatches 0; Indels 0; Gaps 0;

QY 2 VLMAMNII 9
Db :||:|
5 IIMALSII 12

RESULT 37
Q88717 PRELIMINARY; PRT; 24 AA.
AC Q88717;
DT 01-NOV-1998 (TRENBLrel. 08, Created)
DT 01-NOV-1998 (TRENBLrel. 08, Last sequence update)
DT 01-OCT-2002 (TRENBLrel. 22, Last annotation update)
DE Tenascin-r protein (Fragment).
GN TNR OR TENASCIN-R.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RN SEQUENCE FROM N.A.

```

```

RC TISSUE=Brain;
RX MEDLINE=99274730; PubMed=10341229;
RA Weber P., Bartsch U., Rabband M.N., Czaniara R., Lang Y.,
RA Bluetmann H., Margolis R.U., Levinson S.R., Shrager P., Montag D.;
RT "Mice deficient for tenascin-r display alterations of the
RT extracellular matrix and decreased axonal conduction velocities in the
RT CNS.";
RL J. Neurosci. 19:4245-4262(1999).
DR EMBL; AJ005844; CAA06739.1; -.
DR MGd; MGI:99516; Tnr.
FT NON TER 24 24
SQ SEQUENCE 24 AA; 2500 MW; F01852DD5861B23E CRC64;

Query Match 59.5%; Score 25; DB 11; Length 24;
Best Local Similarity 44.4%; Pred. No. 3.1e+02;
Matches 4; Conservative 4; Mismatches 1; Indels 0; Gaps 0;

QY 1 NVLMANNII 9
Db ||:|:|:|
12 NMLIGVNL 20

RESULT 38
Q26413 PRELIMINARY; PRT; 45 AA.
AC Q26413;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-MAR-2002 (TrEMBLrel. 20, Last annotation update)
DE Type II topoisomerase (Fragment).
GN TOP2.
OS Cryptosporidium parvum.
OC Eukaryota; Alveolata; Apicomplexa; Coccidia; Eimeriida;
OC Cryptosporidiidae; Cryptosporidium.
OX NCBI_TaxID=5807;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=95102611; PubMed=7804230;
RA Christopher L.J., Dykstra C.C.;
RT "Identification of a type II topoisomerase gene from Cryptosporidium
RT parvum.";
RL J. Eukaryot. Microbiol. 41:288-288S(1994).
DR EMBL; S76665; AAB32916.2; -.
DR HSSP; P06786; 1BGW.
DR InterPro; IPR002205; DNA_topoisoiV.
DR Pfam; PF00521; DNA_topoisoiV; 1.
DR ProDom; PD000742; DNA_topoisoiV; 1.
KW Isomerase.
FT NON TER 1 1
FT NON TER 45 45
SQ SEQUENCE 45 AA; 4957 MW; B3F0F080C08CFC3F CRC64;

Query Match 59.5%; Score 25; DB 5; Length 45;
Best Local Similarity 55.8%; Pred. No. 5.4e+02;
Matches 5; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 1 NVLMANNII 9
Db ||:|:|:|
35 NVNMAQNFV 43

RESULT 39
Q9K8Y3 PRELIMINARY; PRT; 59 AA.
AC Q9K8Y3;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-MAR-2002 (TrEMBLrel. 15, Last annotation update)
DE Hypothetical protein BH2869.
GN BH2869.
OS Bacillus halodurans.
OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.
OX NCBI_TaxID=86665;

```

```

RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C-125 / JCM 9153;
RX MEDLINE=20512582; PubMed=11058132;
RA Takami H., Nakasone K., Takaki Y., Maeno G., Sasaki R., Masui N.,
RA Fuji F., Hirama C., Nakamura Y., Ogasawara N., Kuhara S.,
RA Horikoshi K.;
RT "Complete genome sequence of the alkaliphilic bacterium Bacillus
RT halodurans and genomic sequence comparison with Bacillus subtilis.";
RL Nucleic Acids Res. 28:4317-4331(2000).
DR EMBL; AP001516; BAB06588.1; -.
KW Hypothetical protein; Complete proteome.
SQ SEQUENCE 59 AA; 7032 MW; 47AFFC9DC9EEF843 CRC64;

Query Match 59.5%; Score 25; DB 16; Length 59;
Best Local Similarity 37.5%; Pred. No. 6.9e+02;
Matches 3; Conservative 5; Mismatches 0; Indels 0; Gaps 0;

QY 2 VLMANNII 9
Db ||:|:|:|
35 ILVALNLV 42

RESULT 40
Q99VZ5 PRELIMINARY; PRT; 74 AA.
AC Q99VZ5;
DT 01-JUN-2001 (TrEMBLrel. 17, Created)
DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT 01-OCT-2002 (TrEMBLrel. 22, Last annotation update)
DE Hypothetical protein SAV0618 (Hypothetical protein MW0582).
GN SAV0618 OR SA0575 OR MW0582.
OS Staphylococcus aureus (strain N315), and
OS Staphylococcus aureus (strain MW2).
OC Bacteria; Firmicutes; Bacillales; Staphylococcus.
OX NCBI_TaxID=158878, 158879, 196620;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=MU50, and N315;
RX MEDLINE=21311952; PubMed=11418146;
RA Kuroda M., Ohta T., Uchiyama I., Baba T., Yuzawa H., Kobayashi I.,
RA Cui L., Oguchi A., Aoki K.-I., Nagai Y., Lian J.-Q., Ito T.,
RA Kanamori M., Matsumaru H., Maruyama A., Murakami H., Hosoyama A.,
RA Mizutani-Ui Y., Takahashi N.K., Sawano T., Inoue R.-I., Kaito C.,
RA Sekimizu K., Hirakawa H., Kuhara S., Goto S., Yabuzaki J.,
RA Kanehisa M., Yamashita A., Oshima K., Furuya K., Yoshino C., Shiba T.,
RA Hattori M., Ogasawara N., Hayashi H., Hiramatsu K.;
RT "Whole genome sequencing of methicillin-resistant Staphylococcus
RT aureus.";
RL Lancet 357:1225-1240(2001).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=MW2;
RX MEDLINE=22040717; PubMed=12044378;
RA Baba T., Takeuchi F., Kuroda M., Yuzawa H., Aoki K.-I., Oguchi A.,
RA Nagai Y., Iwama N., Asano K., Naimi T., Kuroda H., Cui L.,
RA Yamamoto K., Hiramatsu K.;
RT "Genome and virulence determinants of high virulence community-
RT acquired MRSA.";
RL Lancet 359:1819-1827(2002).
DR EMBL; AP003359; BAB56780.1; -.
DR EMBL; AP003131; BAB41807.1; -.
DR EMBL; AP004824; BAB94447.1; -.
KW Hypothetical protein; Complete proteome.
SQ SEQUENCE 74 AA; 8228 MW; 35AFAB2F9FF612AC CRC64;

Query Match 59.5%; Score 25; DB 16; Length 74;
Best Local Similarity 55.6%; Pred. No. 8.4e+02;
Matches 5; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 1 NVLMANNII 9
Db ||:|:|:|

```

```

Db          38 NALVDLNI 46
RESULT 41
Q8U2J3      PRELIMINARY; PRT; 75 AA.
AC Q8U2J3;
DT 01-JUN-2002 (TREMBlrel. 21, Created)
DT 01-JUN-2002 (TREMBlrel. 21, Last sequence update)
DT 01-JUN-2002 (TREMBlrel. 21, Last annotation update)
DE Hypothetical protein PF0841.
GN PF0841.
OS Pyrococcus furiosus.
OC Archaea; Euryarchaeota; Thermococci; Thermococcales; Thermococcaceae;
OC Pyrococcus.
OX NCBI_TaxID=2261;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=VCI / DSM 3638 / ATCC 43587 / JCM 8422;
RA Weiss R.B., Dunn D.M., Robb F.T., Brown J.R.;
RT "The complete sequence of the Pyrococcus furiosus genome."
RL Submitted (FEB-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AE010199; AAL80965.1; -.
KW Hypothetical protein; Complete proteome.
SQ SEQUENCE 75 AA; 8393 MW; A52E0F2EB1852AF1 CRC64;

Query Match 59.5%; Score 25; DB 17; Length 75;
Best Local Similarity 50.5%; Pred. No. 8.5e+02;
Matches 4; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 1 NVLMANNI 8
Db 33 NPLIALNV 40
RESULT 42
Q97P38      PRELIMINARY; PRT; 79 AA.
AC Q97P38;
DT 01-OCT-2001 (TREMBlrel. 18, Created)
DT 01-OCT-2001 (TREMBlrel. 18, Last sequence update)
DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)
DE Conserved domain protein.
GN SPI806.
OS Streptococcus pneumoniae.
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC Streptococcus.
OX NCBI_TaxID=1313;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=TIGR4;
RX MEDLINE=21357209; PubMed=11463916;
RA Tettelin H., Nelson K.E., Paulsen I.T., Eisen J.A., Read T.D.,
RA Peterson S., Heidelberg J., DeBoy R.T., Haft D.H., Dodson R.J.,
RA Durkin A.S., Gwinn M., Kolonay J.P., Nelson W.C., Peterson J.D.,
RA Umayam L.A., White O., Salzberg S.L., Lewis M.R., Radune D.,
RA Holtzapple E., Khouri H., Wolf A.M., Utterback T.R., Hansen C.L.,
RA McDonald L.A., Feldlyum T.V., Angiuoli S., Dickinson T., Hickey E.K.,
RA Holt I.E., Loftus B.J., Yang F., Smith H.O., Venter J.C.,
RA Dougherty B.A., Morrison D.A., Hollingshead S.K., Fraser C.M.;
RT "Complete genome sequence of a virulent isolate of Streptococcus pneumoniae."
RL Science 293:498-506(2001).
DR EMBL; AE007473; AAK75879.1; -.
DR TIGR; SP1806; -.
KW Complete proteome.
SQ SEQUENCE 79 AA; 9470 MW; 9651135ACEB74D09 CRC64;

Query Match 59.5%; Score 25; DB 16; Length 79;
Best Local Similarity 62.5%; Pred. No. 8.9e+02;
Matches 5; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 2 VLMANNII 9
Db 30 ILMINSII 37
RESULT 43
Q8HNB1      PRELIMINARY; PRT; 81 AA.
AC Q8HNB1;
DT 01-MAR-2003 (TREMBlrel. 23, Created)
DT 01-MAR-2003 (TREMBlrel. 23, Last sequence update)
DT 01-MAR-2003 (TREMBlrel. 23, Last annotation update)
DE Cytochrome oxidase subunit II (Fragment).
GN COII.
OS Daktulosphaira vitifoliae (grape phylloxera).
OC Mitochondrion.
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Paraneoptera; Hemiptera; Sternorrhyncha; Aphidiformes;
OC Phylloxeroidea; Phylloxeridae; Daktulosphaira.
OX NCBI_TaxID=58002;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ARIFJARI;
RX MEDLINE=22234187; PubMed=12296945;
RA Downie D.A.;
RT "Locating the sources of an invasive pest, grape phylloxera, using a
RT mitochondrial DNA gene genealogy."
RL Mol. Ecol. 11:2013-2026(2002).
DR EMBL; AF534470; AAN37812.1; -.
KW Mitochondrion.
FT NON_TER 81
SQ SEQUENCE 81 AA; 9706 MW; 87F8E9ED6DEA86E7 CRC64;

Query Match 59.5%; Score 25; DB 8; Length 81;
Best Local Similarity 62.5%; Pred. No. 9.1e+02;
Matches 5; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 2 VLMANNII 9
Db 30 ILMINSII 37
RESULT 44
Q8HNB0      PRELIMINARY; PRT; 81 AA.
AC Q8HNB0;
DT 01-MAR-2003 (TREMBlrel. 23, Created)
DT 01-MAR-2003 (TREMBlrel. 23, Last sequence update)
DT 01-MAR-2003 (TREMBlrel. 23, Last annotation update)
DE Cytochrome oxidase subunit II (Fragment).
GN COII.
OS Daktulosphaira vitifoliae (grape phylloxera).
OC Mitochondrion.
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Paraneoptera; Hemiptera; Sternorrhyncha; Aphidiformes;
OC Phylloxeroidea; Phylloxeridae; Daktulosphaira.
OX NCBI_TaxID=58002;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ARIFAMARI;
RX MEDLINE=22234187; PubMed=12296945;
RA Downie D.A.;
RT "Locating the sources of an invasive pest, grape phylloxera, using a
RT mitochondrial DNA gene genealogy."
RL Mol. Ecol. 11:2013-2026(2002).
DR EMBL; AF534471; AAN37814.1; -.
KW Mitochondrion.
FT NON_TER 81
SQ SEQUENCE 81 AA; 9685 MW; 87F8E9FD6CEA86E7 CRC64;

Query Match 59.5%; Score 25; DB 8; Length 81;
Best Local Similarity 62.5%; Pred. No. 9.1e+02;
Matches 5; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 2 VLMANNII 9
Db 30 ILMINSII 37

```

QY 2 VLMAMNII 9  
:| | | :| | |  
DB 30 ILMIMSII 37

RESULT 45

Q8HNA9 PRELIMINARY; PET; 81 AA.  
AC Q8HNA9;  
DT 01-MAR-2003 (TrEMBLrel. 23, Created)  
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)  
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)  
DE Cytochrome oxidase subunit II (Fragment).  
GN COII.  
OS Daktulosphaira vitifoliae (grape phylloxera).  
OG Mitochondrion.  
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;  
OC Neoptera; Paraneoptera; Hemiptera; Sternorrhyncha; Aphidiformes;  
OC Phylloxeroidea; Phylloxeridae; Daktulosphaira.  
OX NCBI\_TaxID=58002;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=ARIMADARI;  
RX MEDLINE=22234187; PubMed=12296945;  
RA Downie D.A.;  
RT "Locating the sources of an invasive pest, grape phylloxera, using a  
RT mitochondrial DNA gene genealogy.";  
RL Mol. Ecol. 11:2013-2026(2002).  
DR EMBL; AF534472; AA037816.1; -.  
KW Mitochondrion.  
FT NON\_TER 81  
SQ SEQUENCE 81 AA; 9744 MW; 87F8E9ED746B0FA3 CRC64;

Query Match 59.5%; Score 25; DB 8; Length 81;  
Best Local Similarity 62.5%; Pred.No. 9.1e+02;  
Matches 5; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 2 VLMAMNII 9  
:| | | :| | |  
DB 30 ILMIMSII 37

Search completed: February 17, 2004, 10:56:17  
Job time : 14.7426 secs

**THIS PAGE BLANK (USPTO)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: February 17, 2004, 10:50:12 ; Search time 36.1188 Seconds  
(without alignments)  
83.497 Million cell updates/sec

Title: US-09-900-147-3

Perfect score: 91

Sequence: 1 RRVVDALNVLMMNIISK 19

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1107863 seqs, 158726573 residues

Total number of hits satisfying chosen parameters: 717921

Minimum DB seq length: 0

Maximum DB seq length: 100

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : A\_Geneseq\_19Jun03.\*

```
1: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1980.DAT.*
2: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1981.DAT.*
3: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1982.DAT.*
4: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1983.DAT.*
5: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1984.DAT.*
6: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1985.DAT.*
7: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1986.DAT.*
8: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1987.DAT.*
9: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1988.DAT.*
10: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1989.DAT.*
11: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1990.DAT.*
12: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1991.DAT.*
13: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1992.DAT.*
14: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1993.DAT.*
15: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1994.DAT.*
16: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1995.DAT.*
17: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1996.DAT.*
18: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1997.DAT.*
19: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1998.DAT.*
20: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1999.DAT.*
21: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA2000.DAT.*
22: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA2001.DAT.*
23: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA2002.DAT.*
24: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA2003.DAT.*
```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	91	100.0	19	AAW30504	DP-1 transcription
2	91	100.0	37	AAW30501	DP-1 transcription
3	86	94.5	28	AAW57051	E2F activity inhib
4	86	94.5	28	AAW57055	E2F activity inhib
5	83	91.2	19	AAW30515	DP-1 transcription
6	79	86.8	19	AAW30517	DP-1 transcription
7	77	84.6	19	AAW30516	DP-1 transcription
8	76	83.5	16	AAW30506	DP-1 transcription
9	72	79.1	30	AAW30507	DP-1 transcription

10	70	76.9	56	22	AAW67766	Fragment from a wh
11	69	75.8	14	19	AAW30511	DP-1 transcription
12	68	74.7	15	19	AAW57052	E2F activity inhib
13	66	72.5	83	20	AAW32163	Soybean DP-1 prote
14	58	63.7	96	22	ABG58365	Human liver peptid
15	58	63.7	96	22	ABG42955	Peptide #10461 enc
16	58	63.7	96	22	ABW63859	Human bone marrow
17	58	63.7	96	22	AAW76572	Human bone marrow
18	58	63.7	96	22	AAW36777	Peptide #10814 enc
19	58	63.7	96	23	ABG45881	Human peptide enco
20	56	61.5	80	20	AAW32167	Soybean E2F protei
21	54	59.3	11	19	AAW30509	DP-1 transcription
22	51	56.0	20	19	AAW30505	DP-1 transcription
23	51	56.0	85	23	AAU78095	Human DNA binding
24	51	56.0	85	24	ABW82986	Repressor protein
25	50	54.9	76	23	AAU78096	Human DNA binding
26	50	54.9	76	24	ABW82987	Repressor protein
27	49	53.8	28	21	AAW44493	Wheat E2F derived
28	46	50.5	29	19	AAW57054	E2F activity inhib
29	46	50.5	29	19	AAW57049	E2F activity inhib
30	45	49.5	35	23	AAU72601	DEF domain consens
31	42	46.2	9	19	AAW30502	DP-1 transcription
32	40	44.0	24	20	AAW25475	Insulin-like growt
33	40	44.0	25	17	AAW87180	Osteoblast prolife
34	40	44.0	40	17	AAW87181	Osteoblast prolife
35	40	44.0	85	17	AAW87182	Polypeptide sequen
36	39	42.9	80	22	AAU07882	Human prostate spe
37	38	41.8	48	23	ABG66872	Novel human colon
38	38	41.8	63	22	AAU22567	Human digestive sy
39	38	41.8	63	22	AAW92530	Tyrosine activatio
40	37	40.7	24	15	AAW56646	Human immune/haema
41	37	40.7	87	22	AAW89362	E2F activity inhib
42	36	39.6	15	19	AAW57050	Human liver peptid
43	36	39.6	41	22	ABG56095	Peptide #8153 enco
44	36	39.6	41	22	ABW40847	Protein #8503 enco
45	36	39.6	41	22	ABW24904	

#### ALIGNMENTS

RESULT 1  
AAW30504  
ID AAW30504 standard; Peptide; 19 AA.  
XX  
AC AAW30504;  
XX  
DT 26-OCT-1998 (first entry)  
XX  
DE DP-1 transcription factor antagonist peptide H2.  
XX  
DE DP-1; transcription factor; antagonist; E2F protein; apoptosis;  
KW cell proliferation; cardiovascular cell; restenosis; tumour;  
KW surgical stent; therapy.  
XX  
OS Synthetic.  
OS Homo sapiens.  
XX  
FH Key Location/Qualifiers  
FT Peptide 9..18  
FT /note= "Claim 3"  
XX  
PN WO9828334-A1.  
XX  
PD 02-JUL-1998.  
XX  
PF 22-DEC-1997; 97WO-GB03506.  
XX  
PR 20-DEC-1996; 96GB-0026589.  
XX  
PA (PROL-) PROLIFIX LTD.  
XX  
PI Bandara LR, La Thangue NB;

XX WPI; 1998-377596/32.

XX Polypeptide fragments of the DP-1 transcription factor - used for

PT inducing apoptosis, specifically in tumour and cardiovascular cells,

PT e.g. for preventing restenosis

XX Claim 4; Page 44; 55pp; English.

XX Peptide H2 comprises amino acid residues 166-184 in the DEF box

CC (I) (see AAW30501) of transcription factor DP1. Claimed peptides

CC (II) (see AAW30504-07) containing one or both of 2 motifs (see

CC AAW30502-03) of the DEF box are capable of antagonising the

CC heterodimerisation of a DP protein with an E2F protein. Also

CC claimed are variants of these peptides, especially containing

CC substitutions of residues corresponding to residues 167, 169, 171

CC and 175 of DP-1, fusion proteins (III) comprising (I) or (II) and a

CC membrane translocation sequence (see AAW30508), expression vectors

CC encoding (I)-(III) and host cells. (I)-(III) are used

CC therapeutically to induce apoptosis, specifically in tumour or

CC cardiovascular cells, either in vivo or in vitro, e.g. for purging

CC bone marrow. Surgical stents comprising (I)-(III) are used to

CC treat or prevent restenosis in patients who have undergone

CC angioplasty. (I)-(III) function by inactivating the DNA-binding

CC activity of DP/E2F heterodimers. They are also used as research

CC reagents, as positive controls in assays for identifying

CC antagonists of DP-1/E2F dimerisation and as immunoassay agents.

CC Also described is the use of sequences antisense to nucleic acids

CC encoding (I)-(III) to control DP levels in cells, particularly by

CC gene therapy. When formulated with cytotoxic or cytostatic agents,

CC (I)-(III) enhance cell killing.

XX Sequence 19 AA;

XX Query Match 100.0%; Score 91; DB 19; Length 19;

XX Best Local Similarity 100.0%; Pred. No. 2.9e-10;

XX Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 RRRVVDALNVLMMNNIISK 19

Db 1 RRRVVDALNVLMMNNIISK 19

RESULT 2

AAW30501

ID AAW30501 standard; Peptide; 37 AA.

XX AAW30501;

XX 26-OCT-1998 (first entry)

XX DP-1 transcription factor peptide H (DEF box).

XX DP-1; transcription factor; antagonist; E2F protein; apoptosis;

KW cell proliferation; cardiovascular cell; restenosis; tumour;

KW surgical stent; therapy.

XX Synthetic.

OS Homo sapiens.

PN WO9828334-A1.

XX 02-JUL-1998.

XX 22-DEC-1997; 97WO-GB03506.

XX 20-DEC-1996; 96GB-0026589.

XX (PROL-) PROLIFIX LTD.

XX Bandara LR, La Thangue NB;

PI WPI; 1998-377596/32.

XX

XX Polypeptide fragments of the DP-1 transcription factor - used for

PT inducing apoptosis, specifically in tumour and cardiovascular cells,

PT e.g. for preventing restenosis

XX Claim 1; Page 44; 55pp; English.

XX Peptide H (I) comprises residues 163-199, i.e. the DEF box region,

CC of transcription factor DP1. Claimed fragments (II) (see AAW30502-07)

CC of (I) are capable of antagonising the heterodimerisation of a DP

CC protein with an E2F protein. Also claimed are fusion proteins

CC (III) comprising (I) or (II) and a membrane translocation sequence

CC (see AAW30508), expression vectors encoding (I)-(III) and host cells.

CC (I)-(III) are used therapeutically to induce apoptosis,

CC specifically in tumour or cardiovascular cells, either in vivo or in

CC vitro, e.g. for purging bone marrow. Surgical stents comprising

CC (I)-(III) are used to treat or prevent restenosis in patients who

CC have undergone angioplasty. (I)-(III) function by inactivating

CC the DNA-binding activity of DP/E2F heterodimers. They are also

CC used as research reagents, as positive controls in assays for

CC identifying antagonists of DP-1/E2F dimerisation and as immunoassay

CC agents. Also described is the use of sequences antisense to

CC nucleic acids encoding (I)-(III) to control DP levels in cells,

CC particularly by gene therapy. When formulated with cytotoxic

CC or cytostatic agents, (I)-(III) enhance cell killing.

XX Sequence 37 AA;

XX Query Match 100.0%; Score 91; DB 19; Length 37;

XX Best Local Similarity 100.0%; Pred. No. 6.9e-10;

XX Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 RRRVVDALNVLMMNNIISK 19

Db 4 RRRVVDALNVLMMNNIISK 22

RESULT 3

AAW57051

ID AAW57051 standard; peptide; 28 AA.

XX AAW57051;

XX 28-AUG-1998 (first entry)

XX E2F activity inhibiting compound Ib-1.

DE E2F activity; inhibitor; treatment; tumour; arteriosclerosis.

XX Synthetic.

OS Homo sapiens.

XX Key Location/Qualifiers

PH Modified-site 1

FT /note= "N-terminal acetyl"

FT Modified-site 28

FT /note= "C-terminal amide"

XX WO9814474-A1.

XX 09-APR-1998.

XX 26-SEP-1997; 97WO-JP03442.

XX 30-SEP-1996; 96JP-0259432.

XX (KYOW ) KYOWA HAKKO KOGYO KK.

XX Mizukami T, Shibata K, Yamasaki M, Yoshida T;

PI WPI; 1998-240020/21.

XX E2F activity inhibitors - for treatment and prevention of tumours



PT and arteriosclerosis

PS Example 3; Page 27; 52pp; Japanese.

XX This represents a compound that can inhibit E2F activity. The compound is of the formula R1 - A - R2 where R1 is an optionally substituted alkanoyl, allyl, hetero-arylcabonyl, alkoxy carbonyl, aryloxy carbonyl, hetero-aryloxy carbonyl, or H, R2 is OH, or optionally substituted alkoxy or amino, and A is an E2F family dimer forming region or DNA binding region, of at least 12 consecutive amino acids. Compounds of this formula can be used to inhibit E2F activity, and are useful in the treatment and prevention of tumours and arteriosclerosis.

XX SQ Sequence 28 AA;

Query Match 94.5%; Score 86; DB 19; Length 28;

Best Local Similarity 100.0%; Pred. No. 4.2e-09;

Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 RRRVYDALNVLMMNNIIS 18  
 DB 11 RRRVYDALNVLMMNNIIS 28  
 |||||

RESULT 4

AAW57055

ID AAW57055 standard; peptide; 28 AA.

XX AC AAW57055;

XX DT 28-AUG-1998 (first entry)

XX DE E2F activity inhibiting compound IB-3.

XX DE E2F activity; inhibitor; treatment; tumour; arteriosclerosis.

XX OS Synthetic.

XX OS Homo sapiens.

XX Key Location/Qualifiers

FT Modified-site 1 /note= "N-terminal lauroyl"

FT Modified-site 28 /note= "C-terminal amide"

FT WO9814474-A1.

XX PN Mizukami T, Shibata K, Yamasaki M, Yoshida T;

XX PD WPI; 1998-240020/21.

XX PF E2F activity inhibitors - for treatment and prevention of tumours and arteriosclerosis

XX PR Example 7; Page 33; 52pp; Japanese.

XX PS This represents a compound that can inhibit E2F activity. The compound is of the formula R1 - A - R2 where R1 is an optionally substituted alkanoyl, allyl, hetero-arylcabonyl, alkoxy carbonyl, aryloxy carbonyl, hetero-aryloxy carbonyl, or H, R2 is OH, or optionally substituted alkoxy or amino, and A is an E2F family dimer forming region or DNA binding region, of at least 12 consecutive amino acids. Compounds of this formula can be used to inhibit E2F activity, and are useful in the treatment and prevention of tumours and arteriosclerosis.

XX

SQ Sequence 28 AA;

Query Match 94.5%; Score 86; DB 19; Length 28;

Best Local Similarity 100.0%; Pred. No. 4.2e-09;

Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 RRRVYDALNVLMMNNIIS 18  
 DB 11 RRRVYDALNVLMMNNIIS 28  
 |||||

RESULT 5

AAW30515

ID AAW30515 standard; Peptide; 19 AA.

XX AC AAW30515;

XX DT 26-OCT-1998 (first entry)

XX DE DP-1 transcription factor peptide H2mt1.

XX DE DP-1; transcription factor; antagonist; E2F protein; apoptosis; cell proliferation; cardiovascular cell; restenosis; tumour; surgical stent; therapy.

XX OS Synthetic.

XX OS Homo sapiens.

XX Key Location/Qualifiers

FT Misc-difference 4 /note= "V169A mutation"

FT Misc-difference 10 /note= "V175A mutation"

FT WO9828334-A1.

XX PN 02-JUL-1998.

XX PD 22-DEC-1997; 97WO-GB03506.

XX PF 20-DEC-1996; 96GB-0026589.

XX PR (PROL-) PROLIFIX LTD.

XX PA Bandara LR, La Thangue NB;

XX PI WPI; 1998-377596/32.

XX DR Polypeptide fragments of the DP-1 transcription factor - used for inducing apoptosis, specifically in tumour and cardiovascular cells, e.g. for preventing restenosis

XX PS Example D; Page 26; 55pp; English.

XX CC Peptide H2mt1 is based on peptide H2 (see AAW30504) from the DEF box (see AAW30501) of transcription factor DP1. The H2mt1 peptide, in which H2 residues corresponding to DP1 residues Val169 and Val175 are substituted by Ala residues, behaves in a similar fashion to the wild-type H2 peptide in its ability to inactivate E2F site DNA binding activity in D9 EC cell extracts. H2 is an antagonist of the heterodimerisation of DP1 with E2F. Thus, the Val-169 and Val-175 residues of H2 play little role in this activity. H2 and other claimed peptides (see AAW30504-07) from the DEF box region of DP1 can be used to induce apoptosis, specifically in tumour and cardiovascular cells, e.g. for the prevention of restenosis.

XX SQ Sequence 19 AA;

Query Match 91.2%; Score 83; DB 19; Length 19;

Best Local Similarity 89.5%; Pred. No. 9.5e-09;

Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 1 RRRVYDALNVLMMNNIISK 19

Tue Feb 17 11:55:48 2004

us-09-900-147-3-rag

```

Db      1 RRRVDALNVLAMNIISK 19
|||||
RESULT 6
AAW30517
ID AAW30517 standard; Peptide; 19 AA.
XX
AC AAW30517;
XX
XX 26-OCT-1998 (first entry)
XX
DE DP-1 transcription factor peptide H2mt3.
XX
XX DP-1; transcription factor; antagonist; E2F protein; apoptosis;
KW cell proliferation; cardiovascular cell; restenosis; tumour;
KW surgical stent; therapy.
XX
OS Synthetic.
OS Homo sapiens.
XX
XX Key Location/Qualifiers
XX Misc-difference 8 /note= "L173R mutation"
FT
FT Misc-difference 11 /note= "L176R mutation"
FT
XX
XX W09828334-A1.
XX
XX 02-JUL-1998.
XX
XX 22-DEC-1997; 97WO-GB03506.
XX
XX 20-DEC-1996; 96GB-0026589.
XX
XX (PROL-) PROLIFIX LTD.
XX
XX Bandara LR, La Thangue NB;
XX WPI; 1998-377596/32.
XX
XX Polypeptide fragments of the Dp-1 transcription factor - used for
PT inducing apoptosis, specifically in tumour and cardiovascular cells,
PT e.g. for preventing restenosis
XX
XX Example D; Page 26; 55pp; English.
XX
XX Peptide H2mt3 is based on peptide H2 (see AAW30504) from the DEF box
CC (see AAW30501) of transcription factor Dp1. In H2mt3, amino acid
CC residues of H2 that correspond to Dp1 residues Leu173 and Leu176
CC are substituted by Arg residues. H2 is an antagonist of the
CC heterodimerisation of Dp1 with E2F. H2mt3 has none of the
CC antagonistic activity of H2. H2 and other claimed peptides (see
CC AAW30504-07) from the DEF box region of Dp1 can be used to induce
CC apoptosis, specifically in tumour and cardiovascular cells, e.g.
CC for the prevention of restenosis.
XX
XX Sequence 19 AA;
XX
XX Query Match 86.8%; Score 79; DB 19; Length 19;
XX Best Local Similarity 89.5%; Pred. No. 5.4e-08;
XX Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
XX
XX 1 RRRVDALNVLAMNIISK 19
XX |||||
Db      1 RRRVDARNVRLAMNIISK 19
|||||
RESULT 7
AAW30516
ID AAW30516 standard; Peptide; 19 AA.
XX
AC AAW30516;
XX
XX 26-OCT-1998 (first entry)
XX
DE DP-1 transcription factor antagonist peptide H5.
XX
XX DP-1; transcription factor; antagonist; E2F protein; apoptosis;
KW cell proliferation; cardiovascular cell; restenosis; tumour;
KW surgical stent; therapy.
XX
OS Synthetic.
OS Homo sapiens.
XX
XX Key Location/Qualifiers
XX Misc-difference 8 /note= "L173R mutation"
FT
FT Misc-difference 11 /note= "L176R mutation"
FT
XX
XX W09828334-A1.
XX
XX 02-JUL-1998.
XX
XX 22-DEC-1997; 97WO-GB03506.
XX
XX 20-DEC-1996; 96GB-0026589.
XX
XX (PROL-) PROLIFIX LTD.
XX
XX Bandara LR, La Thangue NB;
XX WPI; 1998-377596/32.
XX
XX Polypeptide fragments of the Dp-1 transcription factor - used for
PT inducing apoptosis, specifically in tumour and cardiovascular cells,
PT e.g. for preventing restenosis
XX
XX Example D; Page 26; 55pp; English.
XX
XX Peptide H2mt3 is based on peptide H2 (see AAW30504) from the DEF box
CC (see AAW30501) of transcription factor Dp1. In H2mt3, amino acid
CC residues of H2 that correspond to Dp1 residues Leu173 and Leu176
CC are substituted by Arg residues. H2 is an antagonist of the
CC heterodimerisation of Dp1 with E2F. H2mt3 has none of the
CC antagonistic activity of H2. H2 and other claimed peptides (see
CC AAW30504-07) from the DEF box region of Dp1 can be used to induce
CC apoptosis, specifically in tumour and cardiovascular cells, e.g.
CC for the prevention of restenosis.
XX
XX Sequence 19 AA;
XX
XX Query Match 86.8%; Score 79; DB 19; Length 19;
XX Best Local Similarity 89.5%; Pred. No. 5.4e-08;
XX Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
XX
XX 1 RRRVDALNVLAMNIISK 19
XX |||||
Db      1 RRRVDARNVRLAMNIISK 19
|||||
RESULT 8
AAW30506
ID AAW30506 standard; Peptide; 16 AA.
XX
AC AAW30506;
XX
XX 26-OCT-1998 (first entry)
XX
DE DP-1 transcription factor antagonist peptide H5.
XX
XX DP-1; transcription factor; antagonist; E2F protein; apoptosis;
KW cell proliferation; cardiovascular cell; restenosis; tumour;
KW surgical stent; therapy.
XX
OS Synthetic.
OS Homo sapiens.
XX
XX Key Location/Qualifiers
XX Misc-difference 2 /note= "R167A mutation"
FT
FT Misc-difference 6 /note= "D171A mutation"
FT
XX
XX W09828334-A1.
XX
XX 02-JUL-1998.
XX
XX 22-DEC-1997; 97WO-GB03506.
XX
XX 20-DEC-1996; 96GB-0026589.
XX
XX (PROL-) PROLIFIX LTD.
XX
XX Bandara LR, La Thangue NB;
XX WPI; 1998-377596/32.
XX
XX Polypeptide fragments of the Dp-1 transcription factor - used for
PT inducing apoptosis, specifically in tumour and cardiovascular cells,
PT e.g. for preventing restenosis
XX
XX Example D; Page 26; 55pp; English.
XX
XX Peptide H2mt2 is based on peptide H2 (see AAW30504) from the DEF box
CC (see AAW30501) of transcription factor Dp1. In H2mt2, amino acid
CC residues of H2 that correspond to Dp1 residues Arg167 and Asp171
CC are substituted by Ala residues. H2 is an antagonist of the
CC heterodimerisation of Dp1 with E2F. H2mt2 retains some, but not
CC all, of this antagonistic activity. H2 and other claimed peptides
CC (see AAW30504-07) from the DEF box region of Dp1 can be used to
CC induce apoptosis, specifically in tumour and cardiovascular cells,
CC e.g. for the prevention of restenosis.
XX
XX Sequence 19 AA;
XX
XX Query Match 84.6%; Score 77; DB 19; Length 19;
XX Best Local Similarity 89.5%; Pred. No. 1.3e-07;
XX Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
XX
XX 1 RRRVDALNVLAMNIISK 19
XX |||||
Db      1 RARVYALNVLAMNIISK 19
|||||
```

OS Synthetic.  
 OS Homo sapiens.  
 XX  
 FH Key Location/Qualifiers  
 FT Peptide 5..11  
 FT /note= "Claim 3"  
 FT 7..16  
 FT Peptide /note= "Claim 3"  
 FT  
 PN WO9828334-A1.  
 XX  
 XX  
 PD 02-JUL-1998.  
 XX  
 XX  
 PF 22-DEC-1997; 97WO-GB03506.  
 XX  
 PR 20-DEC-1996; 96GB-0026589.  
 XX  
 PA (PROL-) PROLIFIX LTD.  
 XX  
 PI Bandara LR, La Thangue NB;  
 XX WPI; 1998-377596/32.  
 DR  
 XX  
 XX  
 PT Polypeptide fragments of the DP-1 transcription factor - used for  
 PT inducing apoptosis, specifically in tumour and cardiovascular cells,  
 PT e.g. for preventing restenosis  
 XX  
 PS Claim 4; Page 44; 55pp; English.  
 XX  
 CC Peptide H5 comprises amino acid residues 168-183 in the DEF box  
 CC (I) (see AAW30501) of transcription factor DPI. Claimed peptides  
 CC (II) (see AAW30504-07) containing one or both of 2 motifs (see  
 CC AAW30502-03) of the DEF box are capable of antagonising the  
 CC heterodimerisation of a DP protein with an E2F protein. Also  
 CC claimed are variants of these peptides, especially containing  
 CC substitutions of residues corresponding to residues 167, 169, 171  
 CC and 175 of DP-1, fusion proteins (III) comprising (I) or (II) and a  
 CC membrane translocation sequence (see AAW30508), expression vectors  
 CC encoding (I)-(III) and host cells. (I)-(III) are used  
 CC therapeutically to induce apoptosis, specifically in tumour or  
 CC cardiovascular cells, either in vivo or in vitro, e.g. for purging  
 CC bone marrow. Surgical stents comprising (I)-(III) are used to  
 CC treat or prevent restenosis in patients who have undergone  
 CC angioplasty. (I)-(III) function by inactivating the DNA-binding  
 CC activity of DP/E2F heterodimers. They are also used as research  
 CC reagents, as positive controls in assays for identifying  
 CC antagonists of DP-1/E2F dimerisation and as immunoassay agents.  
 CC Also described is the use of sequences antisense to nucleic acids  
 CC encoding (I)-(III) to control DP levels in cells, particularly by  
 CC gene therapy. When formulated with cytotoxic or cytostatic agents,  
 CC (I)-(III) enhance cell killing.  
 XX  
 SQ Sequence 16 AA;  
 Query Match 83.5%; Score 76; DB 19; Length 16;  
 Best Local Similarity 100.0%; Pred. No. 1.6e-07;  
 Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 3 RYVDALNVLMMNIIS 18  
 DB 1 RYVDALNVLMMNIIS 16  
 RESULT 9  
 AAW30507  
 ID AAW30507 standard; Peptide; 30 AA.  
 XX  
 AC AAW30507;  
 XX  
 DT 26-OCT-1998 (first entry)  
 XX  
 DE DP-1 transcription factor antagonist peptide H7.  
 XX

KW DP-1; transcription factor; antagonist; E2F protein; apoptosis;  
 KW cell proliferation; cardiovascular cell; restenosis; tumour;  
 KW surgical stent; therapy.  
 XX  
 OS Synthetic.  
 OS Homo sapiens.  
 XX  
 FH Key Location/Qualifiers  
 FT Peptide 3..9  
 FT /note= "Claim 3"  
 FT 5..15  
 FT Peptide /note= "Claim 3"  
 FT  
 PN WO9828334-A1.  
 XX  
 XX  
 PD 02-JUL-1998.  
 XX  
 XX  
 PF 22-DEC-1997; 97WO-GB03506.  
 XX  
 PR 20-DEC-1996; 96GB-0026589.  
 XX  
 PA (PROL-) PROLIFIX LTD.  
 XX  
 PI Bandara LR, La Thangue NB;  
 XX WPI; 1998-377596/32.  
 DR  
 XX  
 XX  
 PT Polypeptide fragments of the DP-1 transcription factor - used for  
 PT inducing apoptosis, specifically in tumour and cardiovascular cells,  
 PT e.g. for preventing restenosis  
 XX  
 PS Claim 4; Page 44; 55pp; English.  
 XX  
 CC Peptide H7 comprises amino acid residues 170-199 in the DEF box  
 CC (I) (see AAW30501) of transcription factor DPI. Claimed peptides  
 CC (II) (see AAW30504-07) containing one or both of 2 motifs (see  
 CC AAW30502-03) of the DEF box are capable of antagonising the  
 CC heterodimerisation of a DP protein with an E2F protein. Also  
 CC claimed are variants of these peptides, especially containing  
 CC substitutions of residues corresponding to residues 167, 169, 171  
 CC and 175 of DP-1, fusion proteins (III) comprising (I) or (II) and a  
 CC membrane translocation sequence (see AAW30508), expression vectors  
 CC encoding (I)-(III) and host cells. (I)-(III) are used  
 CC therapeutically to induce apoptosis, specifically in tumour or  
 CC cardiovascular cells, either in vivo or in vitro, e.g. for purging  
 CC bone marrow. Surgical stents comprising (I)-(III) are used to  
 CC treat or prevent restenosis in patients who have undergone  
 CC angioplasty. (I)-(III) function by inactivating the DNA-binding  
 CC activity of DP/E2F heterodimers. They are also used as research  
 CC reagents, as positive controls in assays for identifying  
 CC antagonists of DP-1/E2F dimerisation and as immunoassay agents.  
 CC Also described is the use of sequences antisense to nucleic acids  
 CC encoding (I)-(III) to control DP levels in cells, particularly by  
 CC gene therapy. When formulated with cytotoxic or cytostatic agents,  
 CC (I)-(III) enhance cell killing.  
 XX  
 SQ Sequence 30 AA;  
 Query Match 79.1%; Score 72; DB 19; Length 30;  
 Best Local Similarity 100.0%; Pred. No. 2e-06;  
 Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 5 YDALNVLMMNIISK 19  
 DB 1 YDALNVLMMNIISK 15  
 RESULT 10  
 AAB67766  
 ID AAB67766 standard; peptide; 56 AA.  
 XX  
 AC AAB67766;  
 XX

DT 11-JUN-2001 (first entry)  
XX Fragment from a wheat E2F-dimerisation partner (DP) protein.  
DE E2F-dimerisation partner; DP protein; E2F transcription factor;  
XX GI phase; S phase; cell cycle; retinoblastoma protein;  
KW alter cell proliferation.  
KW  
XX Triticum monococcum.  
XX  
XX WO200121644-A2.  
XX  
XX 29-MAR-2001.  
XX  
XX 25-SEP-2000; 2000WO-EP09325.  
XX  
XX 24-SEP-1999; 99ES-0002127.  
XX  
XX 11-NOV-1999; 99ES-0002474.  
XX  
XX (CNSJ ) CONSEJO SUPERIOR INVESTIGACIONES CIENTIF.  
XX  
XX Gutierrez-Armenta C, Ramirez-Parra E;  
XX  
XX WPI; 2001-257972/26.  
XX  
XX N-PSDB; AAF80148.  
XX  
XX New isolated, enriched, cell free and/or recombinant nucleic acid  
XX useful for e.g. altering cell proliferation characteristic such as to  
XX alter plant cell, organ or tissue size -  
XX  
XX Claim 10; Page 50; 77pp; English.  
XX  
XX AAB67764-68 represent fragments of a wheat E2F-dimerisation partner  
XX (DP) protein. The protein acts as a plant E2F transcription factor.  
XX E2F and DP are two proteins that hetero-dimerise to form an active  
XX transcription factor that regulates G1 to S phase of the cell cycle,  
XX and later, the expression of genes required for S-phase progression.  
XX E2F and retinoblastoma protein also interact as a hetero-dimer in  
XX cells to suppress certain genes. This repression involves binding of  
XX the retinoblastoma protein to the E2F-DP dimer that is in turn bound  
XX to sites on DNA through the E2F DNA binding domain. DP proteins can  
XX be modulated to alter plant cell, organ or tissue shape, and  
XX particularly to alter cell proliferation characteristic such as to  
XX alter plant cell, organ or tissue size.  
XX  
XX Sequence 56 AA;  
Query Match 76.9%; Score 70; DB 22; Length 56;  
Best Local Similarity 68.4%; Pred. No. 1.1e-05;  
Matches 13; Conservative 4; Mismatches 2; Indels 0; Gaps 0;  
QY 1 RRRVYDALNVLMMNIISK 19  
Db 14 RRRVYDAFNVLIALRVIAK 32  
RESULT 11  
AAW30511  
ID AAW30511 standard; Peptide; 14 AA.  
XX  
XX AAW30511;  
XX  
XX 26-OCT-1998 (first entry)  
XX  
XX DP-1 transcription factor peptide H6.  
XX  
XX DP-1; transcription factor; antagonist; E2F protein; apoptosis;  
KW cell proliferation; cardiovascular cell; restenosis; tumour;  
KW surgical stent; therapy.  
XX  
XX Synthetic.  
OS Homo sapiens.  
XX

PN WO9828334-A1.  
XX  
PD 02-JUL-1998.  
XX  
PF 22-DEC-1997; 97WO-GB03506.  
XX  
PR 20-DEC-1996; 96GB-0026589.  
XX  
PA (PROL-) PROLIFIX LTD.  
XX  
PI Bandara LR, La Thangue NB;  
XX  
XX WPI; 1998-377596/32.  
XX  
XX Polypeptide fragments of the DP-1 transcription factor - used for  
XX inducing apoptosis, specifically in tumour and cardiovascular cells,  
XX e.g. for preventing restenosis  
XX  
XX Example C; Page 41; 55pp; English.  
XX  
XX Peptide H6 comprises amino acid residues 167-180 in the DEF box  
XX region (see AAW30501) of transcription factor DP1. Unlike claimed  
XX peptides (see AAW30504-07) that contain one or both of 2 motifs (see  
XX AAW30502-03) of the DP1 DEF box, peptide H6 is not capable of  
XX antagonising the heterodimerisation of a DP protein with an E2F  
XX protein. The claimed peptides, their variants and fusion proteins  
XX can be used to induce apoptosis, specifically in tumour and  
XX cardiovascular cells, e.g. to prevent restenosis.  
XX  
XX Sequence 14 AA;  
Query Match 75.8%; Score 69; DB 19; Length 14;  
Best Local Similarity 100.0%; Pred. No. 2.8e-06;  
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 2 RRVYDALNVLMMN 15  
Db 1 RRVYDALNVLMMN 14  
RESULT 12  
AAW57052  
ID AAW57052 standard; peptide; 15 AA.  
XX  
XX AAW57052;  
XX  
XX 28-AUG-1998 (first entry)  
XX  
XX E2F activity inhibiting compound Ib-2.  
DE  
XX  
XX E2F activity; inhibitor; treatment; tumour; arteriosclerosis.  
KW  
XX  
XX Synthetic.  
OS Homo sapiens.  
XX  
XX Key Location/Qualifiers  
XX Modified-site 1 /note= "N-terminal acetyl"  
XX Modified-site 15  
XX /note= "C-terminal amide"  
XX  
XX WO9814474-A1.  
XX  
XX 09-APR-1998.  
XX  
XX 26-SEP-1997; 97WO-JF03442.  
XX  
XX 30-SEP-1996; 96JP-0259432.  
XX  
XX (KYOW ) KYOWA HAKKO KOGYO KK.  
XX  
XX Mizukami T, Shibata K, Yamasaki M, Yoshida T;  
PI  
XX

DR WPI; 1998-240020/21.  
XX E2F activity inhibitors - for treatment and prevention of tumours  
PT and arteriosclerosis  
XX  
XX Example 4; Page 28; 52pp; Japanese.  
XX  
XX This represents a compound that can inhibit E2F activity. The compound  
CC is of the formula R1 - A - R2 where R1 is an optionally substituted  
CC alkanoyl, allyl, hetero-arylcarbonyl, alkoxycarbonyl, aryloxy carbonyl,  
CC hetero-aryloxy carbonyl, or H. R2 is OH, or optionally substituted alkoxy  
CC or amino, and A is an E2F family dimer forming region or DNA binding  
CC region, of at least 12 consecutive amino acids. Compounds of this formula  
CC can be used to inhibit E2F activity, and are useful in the treatment and  
CC prevention of tumours and arteriosclerosis.  
XX  
XX Sequence 15 AA;  
SQ  
Query Match 74.7%; Score 68; DB 19; Length 15;  
Best Local Similarity 100.0%; Pred. No. 4.8e-06;  
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 RRRVYDALNVLMMAM 14  
DB 2 RRRVYDALNVLMMAM 15  
RESULT 13  
AAY32163  
ID AAY32163 standard; Protein; 83 AA.  
XX  
AC AAY32163;  
XX  
XX 01-FEB-2000 (first entry)  
DT  
DE Soybean DP-1 protein fragment.  
XX  
XX DP-1; soybean; cell cycle regulatory protein;  
KW transcription factor; herbicide.  
XX  
XX Glycine max.  
XX  
FH Key Location/Qualifiers  
FT Misc-difference 10 /note= "encoded by GNC"  
FT Misc-difference 25 /note= "encoded by ANT"  
FT Misc-difference 26 /note= "encoded by GNC"  
FT Misc-difference 34 /note= "encoded by GGN"  
FT Misc-difference 35 /note= "encoded by CNT"  
FT Misc-difference 49 /note= "encoded by NAN"  
FT Misc-difference 54 /note= "encoded by ANG"  
FT Misc-difference 55 /note= "encoded by NAT"  
FT Misc-difference 59 /note= "encoded by NAG"  
FT Misc-difference 63 /note= "encoded by NAT"  
FT Misc-difference 69 /note= "encoded by GNG"  
FT Misc-difference 71 /note= "encoded by NTC"  
FT Misc-difference 80 /note= "encoded by CNA"  
XX WO953075-A2.  
XX 21-OCT-1999.  
PD

XX  
XX 08-APR-1999; 99WO-US07638.  
XX  
XX 09-APR-1998; 98US-0081132.  
XX  
XX (DUPO ) DU PONT DE NEMOURS & CO E I.  
XX  
XX Klein TM, Morakinyo LO, Odell JT, Sakai H;  
PI WPI; 1999-633830/54.  
XX N-PSDB; AA234579.  
DR  
XX Plant-derived cell cycle regulatory proteins -  
XX  
XX Claim 10; Page 41; 44pp; English.  
XX  
XX This sequence represents 42% of the middle region of soybean cell  
CC cycle regulatory protein DP-1, as deduced from an isolated  
CC cDNA clone (see AA234579). The invention relates to nucleic acid  
CC fragments (see AA234575-83) encoding plant CDC-16, DP-1, DP-2 and  
CC E2F cell cycle regulatory proteins (see AAY32159-67). It also  
CC relates to the construction of a chimeric gene encoding all or a  
CC portion of the cell cycle regulatory protein, in sense or antisense  
CC orientation, where expression of the chimeric gene results in  
CC production of altered levels of the cell cycle regulatory protein in  
CC a transformed host cell. The nucleic acids and proteins may be  
CC used to facilitate studies of cell cycle regulation in plants,  
CC provide genetic tools to enhance cell growth in tissue culture,  
CC increase gene transfer efficiency and provide more stable  
CC transformations. The proteins may also provide targets to  
CC facilitate design and/or identification of cell cycle regulatory  
CC proteins that may be useful as herbicides.  
XX  
XX Sequence 83 AA;  
SQ  
Query Match 72.5%; Score 66; DB 20; Length 83;  
Best Local Similarity 83.3%; Pred. No. 0.0001;  
Matches 15; Conservative 1; Mismatches 2; Indels 0; Gaps 0;  
QY 1 RRRVYDALNVLMMAMNIIS 18  
DB 7 RRRXDDALNVLMMAMDIIS 24  
RESULT 14  
ABG58365  
ID ABG58365 standard; Peptide; 96 AA.  
XX  
AC ABG58365;  
XX  
DT 25-FEB-2003 (first entry)  
XX  
DE Human liver peptide, SEQ ID No 37013.  
XX  
XX Human; liver; cirrhosis; hyperlipoproteinaemia; hyperlipidaemia;  
KW hypercholesterolaemia; coronary heart disease.  
XX  
XX Homo sapiens.  
XX  
XX WO200157273-A2.  
XX  
XX 09-AUG-2001.  
XX  
XX 30-JAN-2001; 2001WO-US000664.  
XX  
XX 04-FEB-2000; 2000US-0180312.  
PR 26-MAY-2000; 2000US-0207456.  
PR 30-JUN-2000; 2000US-0608408.  
PR 03-AUG-2000; 2000US-0632366.  
PR 21-SEP-2000; 2000US-0234687.  
PR 27-SEP-2000; 2000US-0236359.  
PR 04-OCT-2000; 2000GB-0024263.  
XX

PA (MOLE-) MOLECULAR DYNAMICS INC.  
 XX Penn SG, Hanzel DK, Chen W, Rank DR;  
 XX WPI; 2001-488898/53.  
 XX Human genome-derived single exon nucleic acid probes useful for  
 PT analysing gene expression in human adult liver -  
 XX Claim 27; SEQ ID No 37013; 658pp; English.  
 XX The invention relates to a single exon nucleic acid probe (SENP) (I) for  
 CC measuring human gene expression in a sample derived from human adult  
 CC liver, comprising one of 13109 defined nucleotide sequences given in the  
 CC specification (or complements/ fragments). The probe hybridises at high  
 CC stringency to a nucleic acid molecule expressed in the human adult  
 CC liver. (I) may be used for predicting, measuring and displaying gene  
 CC expression in samples derived from human adult liver. The genes  
 CC identified may be involved in genetic liver diseases such as cirrhosis,  
 CC hyperlipoproteinaemia, hyperlipidaemia and hypercholesterolaemia which  
 CC is associated with coronary heart disease. ABG47348-ABG59930 represent  
 CC human liver single exon encoded peptides of the invention.  
 CC Note: The sequence information for this patent does not appear in the  
 CC printed specification but was obtained in electronic format directly  
 CC from WIPO at ftp.wipo.int/pub/published\_pct\_sequences.  
 XX SQ Sequence 96 AA;  
 Query Match 63.7%; Score 58; DB 22; Length 96;  
 Best Local Similarity 47.4%; Pred. No. 0.004;  
 Matches 9; Conservative 8; Mismatches 2; Indels 0; Gaps 0;  
 OY 1 RRRVVDALNVLMMNIISK 19  
 DB 3 RRRYDIVNVLESLHLVSR 21  
 RESULT 15  
 ABB42955  
 ID ABB42955 standard; Peptide; 96 AA.  
 XX ABB42955;  
 XX 04-FEB-2002 (first entry)  
 DE Peptide #10461 encoded by human foetal liver single exon probe.  
 XX Human; foetal liver; gene expression; single exon nucleic acid probe.  
 XX Homo sapiens.  
 XX WO200157277-A2.  
 XX 09-AUG-2001.  
 XX 30-JAN-2001; 2001WO-US00669.  
 XX 04-FEB-2000; 2000US-0180312.  
 XX 26-MAY-2000; 2000US-0207456.  
 XX 30-JUN-2000; 2000US-0608408.  
 XX 03-AUG-2000; 2000US-0632366.  
 XX 21-SEP-2000; 2000US-0234687.  
 XX 27-SEP-2000; 2000US-0236359.  
 XX 04-OCT-2000; 2000GB-0024263.  
 XX (MOLE-) MOLECULAR DYNAMICS INC.  
 XX Penn SG, Hanzel DK, Chen W, Rank DR;  
 XX WPI; 2001-483447/52.  
 XX Human genome-derived single exon nucleic acid probes useful for  
 PT analysing gene expression in human fetal liver -

XX PS Claim 27; SEQ ID NO 35590; 639pp + sequence listing; English.  
 XX The invention relates to a single exon nucleic acid probe for  
 CC measuring human gene expression in a sample derived from human foetal  
 CC liver. The single exon nucleic acid probes may be used for predicting,  
 CC measuring and displaying gene expression in samples derived from human  
 CC foetal liver. The present sequence is a peptide encoded by a single exon  
 CC nucleic acid probe of the invention.  
 CC Note: The sequence data for this patent did not form part of the  
 CC printed specification, but was obtained in electronic format directly  
 CC from WIPO at ftp.wipo.int/pub/published\_pct\_sequences.  
 XX SQ Sequence 96 AA;  
 Query Match 63.7%; Score 58; DB 22; Length 96;  
 Best Local Similarity 47.4%; Pred. No. 0.004;  
 Matches 9; Conservative 8; Mismatches 2; Indels 0; Gaps 0;  
 OY 1 RRRVVDALNVLMMNIISK 19  
 DB 3 RRRYDIVNVLESLHLVSR 21  
 RESULT 16  
 AAM63859  
 ID AAM63859 standard; Protein; 96 AA.  
 XX AAM63859;  
 XX 05-NOV-2001 (first entry)  
 DT Human brain expressed single exon probe encoded protein SEQ ID NO: 35964.  
 XX Human; brain expressed exon; gene expression analysis; probe;  
 KW microarray; Alzheimer's disease; multiple sclerosis; schizophrenia;  
 KW epilepsy; cancer.  
 XX Homo sapiens.  
 XX WO200157275-A2.  
 XX 09-AUG-2001.  
 XX 30-JAN-2001; 2001WO-US00667.  
 XX 04-FEB-2000; 2000US-0180312.  
 XX 26-MAY-2000; 2000US-0207456.  
 XX 30-JUN-2000; 2000US-0608408.  
 XX 03-AUG-2000; 2000US-0632366.  
 XX 21-SEP-2000; 2000US-0234687.  
 XX 27-SEP-2000; 2000US-0236359.  
 XX 04-OCT-2000; 2000GB-0024263.  
 XX (MOLE-) MOLECULAR DYNAMICS INC.  
 XX Penn SG, Hanzel DK, Chen W, Rank DR;  
 XX WPI; 2001-483446/52.  
 XX Single exon nucleic acid probes for analyzing gene expression in human  
 XX brains -  
 XX Example 4; SEQ ID NO: 35964; 650pp + Sequence Listing; English.  
 XX The present invention provides a number of single exon nucleic acid  
 CC probes which are derived from genomic sequences expressed in the human  
 CC brain. They can be used to measure gene expression in brain cell samples,  
 CC which may enable the diagnosis and improved treatment of nervous system  
 CC diseases such as Alzheimer's disease, multiple sclerosis, schizophrenia,  
 CC epilepsy and cancers. The present sequence is a protein encoded by one of  
 CC the probes of the invention.

SQ Sequence 96 AA;  
Query Match 63.7%; Score 58; DB 22; Length 96;  
Best Local Similarity 47.4%; Pred. No. 0.004; Mismatches 2; Indels 0; Gaps 0;  
Matches 9; Conservative 8;  
QY 1 RRRVYDALNVLMAMNIISK 19  
DB 3 RRRYDIVNVLESILVSR 21  
RESULT 17  
AAM76672  
ID AAM76672 standard; Protein; 96 AA.  
XX AC AAM76672;  
XX DT 06-NOV-2001 (first entry)  
XX DE Human bone marrow expressed probe encoded protein SEQ ID NO: 36978.  
XX KW Human; bone marrow expressed exon; gene expression analysis; probe;  
XX KW microarray; cancer; leukaemia; lymphoma; myeloma.  
XX OS Homo sapiens.  
XX PN WO200157276-A2.  
XX PD 09-AUG-2001.  
XX PF 30-JAN-2001; 2001WO-US00668.  
XX PR 04-FEB-2000; 2000US-0180312.  
XX PR 26-MAY-2000; 2000US-0207456.  
XX PR 30-JUN-2000; 2000US-0608408.  
XX PR 03-AUG-2000; 2000US-0632366.  
XX PR 21-SEP-2000; 2000US-0234687.  
XX PR 27-SEP-2000; 2000US-0236359.  
XX PR 04-OCT-2000; 2000GB-0024263.  
XX PA (MOLE-) MOLECULAR DYNAMICS INC.  
XX PI Penn SG, Hanzel DK, Chen W, Rank DR;  
XX PD WPI; 2001-488900/53.  
XX PF Human genome-derived single exon nucleic acid probes useful for  
XX PT analyzing gene expression in human bone marrow -  
XX PS Example 4; SEQ ID NO: 36978; 659pp + Sequence Listing; English.  
XX CC The present invention provides a number of single exon nucleic acid  
XX CC probes which are derived from genomic sequences expressed in the human  
XX CC bone marrow. They can be used to measure gene expression in bone marrow  
XX CC samples, which may enable the improved diagnosis and treatment of cancers  
XX CC such as lymphoma, leukaemia and myeloma. The present sequence is a  
XX CC protein encoded by one of the probes of the invention.  
XX SQ Sequence 96 AA;  
Query Match 63.7%; Score 58; DB 22; Length 96;  
Best Local Similarity 47.4%; Pred. No. 0.004; Mismatches 2; Indels 0; Gaps 0;  
Matches 9; Conservative 8;  
QY 1 RRRVYDALNVLMAMNIISK 19  
DB 3 RRRYDIVNVLESILVSR 21  
RESULT 18  
AAM36777  
ID AAM36777 standard; Protein; 96 AA.  
XX

AC AAM36777;  
XX DT 17-OCT-2001 (first entry)  
XX DE Peptide #10814 encoded by probe for measuring placental gene expression.  
XX KW Probe; microarray; human; placenta; antenatal diagnosis;  
XX KW genetic disorder.  
XX OS Homo sapiens.  
XX PN WO200157272-A2.  
XX PD 09-AUG-2001.  
XX PF 30-JAN-2001; 2001WO-US00663.  
XX PR 04-FEB-2000; 2000US-0180312.  
XX PR 26-MAY-2000; 2000US-0207456.  
XX PR 30-JUN-2000; 2000US-0608408.  
XX PR 03-AUG-2000; 2000US-0632366.  
XX PR 21-SEP-2000; 2000US-0234687.  
XX PR 27-SEP-2000; 2000US-0236359.  
XX PR 04-OCT-2000; 2000GB-0024263.  
XX PA (MOLE-) MOLECULAR DYNAMICS INC.  
XX PI Penn SG, Hanzel DK, Chen W, Rank DR;  
XX PD WPI; 2001-488997/53.  
XX PF Human genome-derived single exon nucleic acid probes useful for  
XX PT analyzing gene expression in human placenta -  
XX PS Claim 27; SEQ ID No 37046; 654pp; English.  
XX CC The present invention relates to single exon nucleic acid probes (SENP:  
XX CC see AAI31315-AAI57546). The present sequence is a peptide encoded by one  
XX CC such probe. The probes are useful for producing a microarray for  
XX CC predicting, measuring and displaying gene expression in samples derived  
XX CC from human placenta. The probes are useful for antenatal diagnosis of  
XX CC human genetic disorders.  
XX SQ Sequence 96 AA;  
Query Match 63.7%; Score 58; DB 22; Length 96;  
Best Local Similarity 47.4%; Pred. No. 0.004; Mismatches 2; Indels 0; Gaps 0;  
Matches 9; Conservative 8;  
QY 1 RRRVYDALNVLMAMNIISK 19  
DB 3 RRRYDIVNVLESILVSR 21  
RESULT 19  
ABG45881  
ID ABG45881 standard; Peptide; 96 AA.  
XX AC ABG45881;  
XX DT 19-AUG-2002 (first entry)  
XX DE Human peptide encoded by genome-derived single exon probe SEQ ID 35546.  
XX KW Human; single exon probe; asthma; lung cancer; COPD; ILD;  
XX KW chronic obstructive pulmonary disease; interstitial lung disease;  
XX KW familial idiopathic pulmonary fibrosis; neurofibromatosis;  
XX KW tuberous sclerosis; Gaucher's disease; Niemann-Pick disease;  
XX KW Hermansky-Pudlak syndrome; sarcoidosis; pulmonary haemosiderosis;  
XX KW pulmonary histiocytosis; lymphangioma; lymphocystic pulmonary dysplasia;  
XX KW pulmonary alveolar proteinosis; fibrocystic pulmonary hypertension;  
XX KW primary ciliary dyskinesia; pulmonary hypertension;  
XX KW hyaline membrane disease.

Matches 9; Conservative 8; Mismatches 2; Indels 0; Gaps 0;

QY 1 RRRYDALNVLMMNIISK 19  
 DB 3 RRRYDIVNVLESILVSR 21

RESULT 20  
 AAY32167  
 ID AAY32167 standard; Protein; 80 AA.  
 AC AAY32167;  
 DT 01-FEB-2000 (first entry)  
 DE Soybean E2F protein fragment.  
 KW E2F; soybean; cell cycle regulatory protein; transcription factor;  
 herbicide.  
 OS Glycine max.  
 PN WO9953075-A2.  
 PD 21-OCT-1999.  
 PF 08-APR-1999; 99WO-US07638.  
 PR 09-APR-1998; 98US-0081132.  
 PA (DUPO) DU PONT DE NEMOURS & CO E I.  
 PI Klein TW, Morakinyo LO, Odell JT, Sakai H;  
 WPI; 1999-633830/54.  
 DR N-PSDB; AA234583.  
 PT Plant-derived cell cycle regulatory proteins -  
 PS Claim 20; Page 44-45; 44pp; English.

(MOLE-) MOLECULAR DYNAMICS INC.

Penn SG, Hanzel DK, Chen W, Rank DR;  
 WPI; 2002-114183/15.  
 Spatially-addressable set of single exon nucleic acid probes, used to  
 measure gene expression in human lung samples -  
 Claim 27; SEQ ID No 35546; 634pp; English.

The invention relates to a spatially-addressable set of single exon  
 nucleic acid probes for measuring gene expression in a sample derived  
 from human lung comprising single exon nucleic acid probes having one of  
 12614 nucleic acid sequences mentioned in the specification, or their  
 complements or the 12387 open reading frames derived from the 12614  
 probes. Also included are a microarray comprising the novel set of  
 probes; the novel set of probes which hybridize at high stringency to a  
 nucleic acid expressed in the human lung; measuring gene expression in a  
 sample derived from human lung, comprising (a) contacting the array with  
 a collection of detectably labeled nucleic acids derived from human lung  
 mRNA, and (b) measuring the label detectably bound to each probe of  
 the array; identifying exons in a eukaryotic genome, comprising  
 (a) algorithmically predicting at least one exon from genomic sequences  
 of the eukaryote; and (b) detecting specific hybridization of detectably  
 labeled nucleic acids from eukaryote lung mRNA, to a single exon probe,  
 having a fragment identical to the predicted exon, the probe is included  
 in the above mentioned microarray; assigning exons to a single gene,  
 comprising (a) identifying exons from genomic sequence by the method  
 above and (b) measuring the expression of each of the exons in several  
 tissues and/or cell types using hybridization to a single exon  
 microarrays having a probe with the exon, where a common pattern of  
 expression of the exons in the tissues and/or cell types indicates that  
 the exons should be assigned to a single gene; a peptide comprising one  
 of 12011 sequences, mentioned in the specification, or encoded by the  
 probes/open reading frames (ORF). The probes are used for gene  
 expression analysis, and for identifying exons in a gene, particularly  
 using human lung derived mRNA and for the study of lung diseases  
 such as asthma, lung cancer, chronic obstructive pulmonary disease  
 (COPD), interstitial lung disease (ILD), familial idiopathic pulmonary  
 fibrosis, neurofibromatosis, tuberous sclerosis, Gaucher's disease,  
 Niemann-Pick disease, Hermansky-Pudlak syndrome, lymphangioleiomyomatosis,  
 haemosiderosis, pulmonary histiocytosis, Karagenen syndrome, fibrocystic  
 pulmonary alveolar proteinosis, primary ciliary dyskinesia, pulmonary hypertension  
 and hyaline membrane disease. The present sequence is a peptide/protein  
 encoded by a single exon probe of the invention.  
 Note: The sequence data for this patent did not form part  
 of the printed specification, but was obtained in electronic  
 format directly from WIPO at  
 ftp.wipo.int/pub/published\_pct\_sequences.

Sequence 96 AA;

Query Match 63.7%; Score 58; DB 23; Length 96;  
 Best Local Similarity 47.4%; Pred. No. 0.004;

Query Match 61.5%; Score 56; DB 20; Length 80;  
 Best Local Similarity 47.4%; Pred. No. 0.0076;  
 Matches 9; Conservative 7; Mismatches 3; Indels 0; Gaps 0;

QY 1 RRRYDALNVLMMNIISK 19  
 DB 53 RRRYDIVNVLESIGVLSR 71

RESULT 21  
 AAW30509  
 ID AAW30509 standard; Peptide; 11 AA.  
 XX



```

AC AAW30509;
DT 26-OCT-1998 (first entry)
DE DP-1 transcription factor peptide H1.
XX DP-1; transcription factor; antagonist; E2F protein; apoptosis;
KW cell proliferation; cardiovascular cell; restenosis; tumour;
KW surgical stent; therapy.
XX Synthetic.
OS Homo sapiens.
XX WO9828334-A1.
PN 02-JUL-1998.
XX 22-DEC-1997; 97WO-GB03506.
PF 20-DEC-1996; 96GB-0026589.
XX (PROL-) PROLIFIX LTD.
XX Bandara LR, La Thangue NB;
PI WPI; 1998-377596/32.
XX Polypeptide fragments of the DP-1 transcription factor - used for
PT inducing apoptosis, specifically in tumour and cardiovascular cells,
PT e.g. for preventing restenosis
XX Example C; Page 41; 55pp; English.
XX Peptide H1 comprises amino acid residues 166-176 in the DEF box
CC region (see AAW30501) of transcription factor DPI. Unlike claimed
CC peptides (see AAW30504-07) that contain one or both of 2 motifs (see
CC AAW30502-03) of the DPI DEF box, peptide H1 is not capable of
CC antagonising the heterodimerisation of a DP protein with an E2F
CC protein. The claimed peptides, their variants and fusion proteins
CC can be used to induce apoptosis, specifically in tumour and
CC cardiovascular cells, e.g. to prevent restenosis.
XX Sequence 11 AA;
SQ Query Match 59.3%; Score 54; DB 19; Length 11;
Best Local Similarity 100.0%; Pred. No. 0.0014;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 RRRVYDALNVL 11
Db 1 RRRVYDALNVL 11
RESULT 22
AAW30505
ID AAW30505 standard; Peptide; 20 AA.
XX AAW30505;
AC AAW30505;
DT 26-OCT-1998 (first entry)
XX DP-1 transcription factor antagonist peptide H3.
DE DP-1; transcription factor; antagonist; E2F protein; apoptosis;
KW cell proliferation; cardiovascular cell; restenosis; tumour;
KW surgical stent; therapy.
XX Synthetic.
OS Homo sapiens.
XX Key Location/Qualifiers
FH Peptide 1..10
FT /note= "Claim 3"

```

```

XX WO9828334-A1.
PN 02-JUL-1998.
XX 22-DEC-1997; 97WO-GB03506.
PF 20-DEC-1996; 96GB-0026589.
XX (PROL-) PROLIFIX LTD.
XX Bandara LR, La Thangue NB;
PI WPI; 1998-377596/32.
XX Polypeptide fragments of the DP-1 transcription factor - used for
PT inducing apoptosis, specifically in tumour and cardiovascular cells,
PT e.g. for preventing restenosis
XX Claim 4; Page 44; 55pp; English.
XX Peptide H3 comprises amino acid residues 174-193 in the DEF box
CC (I) (see AAW30501) of transcription factor DPI. Claimed peptides
CC (II) (see AAW30504-07) containing one or both of 2 motifs (see
CC AAW30502-03) of the DEF box are capable of antagonising the
CC heterodimerisation of a DP protein with an E2F protein. Also
CC claimed are variants of these peptides, especially containing
CC substitutions of residues corresponding to residues 167, 169, 171
CC and 175 of DP-1, fusion proteins (III) comprising (I) or (II) and a
CC membrane translocation sequence (see AAW30508), expression vectors
CC encoding (I)-(III) and host cells. (I)-(III) are used
CC therapeutically to induce apoptosis, specifically in tumour or
CC cardiovascular cells, either in vivo or in vitro, e.g. for purging
CC bone marrow. Surgical stents comprising (I)-(III) are used to
CC treat or prevent restenosis in patients who have undergone
CC angioplasty. (I)-(III) function by inactivating the DNA-binding
CC activity of DP/E2F heterodimers. They are also used as research
CC reagents, as positive controls in assays for identifying
CC antagonists of DP-1/E2F dimerisation and as immunoassay agents.
CC Also described is the use of sequences antisense to nucleic acids
CC encoding (I)-(III) to control DP levels in cells, particularly by
CC gene therapy. When formulated with cytotoxic or cytostatic agents,
CC (I)-(III) enhance cell killing.
XX Sequence 20 AA;
SQ Query Match 56.0%; Score 51; DB 19; Length 20;
Best Local Similarity 100.0%; Pred. No. 0.011;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 9 NVLMAMNIISK 19
Db 1 NVLMAMNIISK 11
RESULT 23
AAU78095
ID AAU78095 standard; Protein; 85 AA.
XX AAU78095;
AC AAU78095;
DT 18-JUN-2002 (first entry)
XX Human DNA binding domain E2F-1.
DE Human; telomerase reverse transcriptase; TERT; Site C; Progeria; burn;
KW repressor binding site; Hutchinson-Gilford syndrome; AIDS; cancer;
KW acquired immunodeficiency syndrome; cardiovascular disease; osteoporosis;
KW skin rejuvenation; immune senescence; bone marrow transplant; skin graft;
KW neoplastic disease; TERT minimal promoter; DNA binding domain; E2F-1.
XX Homo sapiens.
OS Homo sapiens.
XX

```

FW Key Location/Qualifiers  
FT Domain 18..59  
FT /label= DNA\_binding\_domain  
XX WO200216657-A1.  
XX 28-FEB-2002.  
XX 17-AUG-2001; 2001WO-US25861.  
XX 24-AUG-2000; 2000US-227865P.  
PR 01-SEP-2000; 2000US-230174P.  
PR 05-OCT-2000; 2000US-238345P.  
XX (SIER-) SIERRA SCI INC.  
XX Andrews WH, Foster CA, Fraser S, Mohammadpour H;  
XX WPI; 2002-280952/32.  
XX Modulating expression of telomerase reverse transcriptase (TERT) in a  
PT cell, for regulating proliferative capacity of a cell, involves  
PT modulating TERT transcription repression by Site C repressor binding  
PT site -  
XX Disclosure; Page 8; 66pp; English.  
XX The present invention relates to a new method of modulating expression  
XX of telomerase reverse transcriptase (TERT) from a TERT expression system  
XX that includes a TERT promoter and a Site C repressor binding site. The  
XX method of the invention involves modulating TERT transcription repression  
XX by the Site C repressor binding site. The method of the invention is  
XX useful for modulating expression of TERT for producing a mammalian  
XX antibody. The method is also useful in a variety of different  
XX applications, including immortalisation of cells, production of reagents  
XX for use in life science research, therapeutic applications, and  
XX therapeutic agent screening applications. Increasing TERT expression  
XX delays natural telomeric shortening and/or increases telomeric length and  
XX is useful for treating disease conditions such as progeria or  
XX Hutchinson-Gilford syndrome, acquired immunodeficiency syndrome (AIDS),  
XX cardiovascular disease, osteoporosis, in skin rejuvenation and to inhibit  
XX immune senescence. The method can be employed to lengthen telomeres of  
XX osteoblast and osteoclast stem cells, encouraging bone replacement and  
XX proper remodeling and reinforcement, and can thus be used in bone marrow  
XX transplants for the treatment of cancer and skin grafts for burn  
XX victims and as such the method improves the survival and effectiveness of  
XX bone marrow and skin cell transplants. Decreasing TERT expression is  
XX useful for treating cellular proliferative disease conditions, including  
XX neoplastic disease conditions e.g. cancer. The present amino acid  
XX sequence represents the human DNA binding domain E2F-1.  
XX Sequence 85 AA;  
Query Match 56.0%; Score 51; DB 23; Length 85;  
Best Local Similarity 47.4%; Pred. No. 0.072;  
Matches 9; Conservative 5; Mismatches 5; Indels 0; Gaps 0;  
QY 1 RRRVVDALNVLMMNIISK 19  
Db :||:| | | | | :||:|  
55 KRRIVDTNVLEGIQLIAK 73  
RESULT 24  
ABB82986  
ID ABB82986 standard; Protein; 85 AA.  
XX ABB82986;  
XX 14-APR-2003 (first entry)  
DE Repressor protein E2F-1 DNA binding domain.  
XX Telomerase reverse transcriptase; TERT; Site C repressor; transcription;

KW cytostatic; immunostimulant; anti-HIV; vulneryary; telomerase; human;  
KW repressor protein; E2F-1.  
XX Homo sapiens.  
XX WO2002101010-A2.  
XX 19-DEC-2002.  
XX 06-JUN-2002; 2002WO-US17959.  
XX 07-JUN-2001; 2001US-296992P.  
XX (SIER-) SIERRA SCI INC.  
XX Foster CA, Fraser S, Mohammadpour H, Andrews WH;  
XX WPI; 2003-167401/16.  
XX Modulating expression of telomerase reverse transcriptase TERT by  
PT blocking repression of TERT transcription, useful for the diagnosis and  
PT treatment of disorders associated with aberrant telomerase activity  
PT such as cancer and HIV  
XX Disclosure; Page 7; 47pp; English.  
XX The invention relates to modulating expression of telomerase reverse  
XX transcriptase (TERT) expression system that includes a TERT promoter and  
XX a Site C repressor binding site. The method involves modulating TERT  
XX transcription repression by the Site C repressor binding site. The  
XX methods and compositions of the present invention are useful for the  
XX immortalization of cells, production of reagents in life science  
XX research, therapeutic agent screening applications, diagnosis and  
XX treatment of disorders associated with aberrant telomerase activity such  
XX as cancer, progeria, immune senescence, HIV, and in skin rejuvenation.  
XX The present sequence represents the DNA binding domain of a repressor  
XX protein E2F-1.  
XX Sequence 85 AA;  
Query Match 56.0%; Score 51; DB 24; Length 85;  
Best Local Similarity 47.4%; Pred. No. 0.072;  
Matches 9; Conservative 5; Mismatches 5; Indels 0; Gaps 0;  
QY 1 RRRVVDALNVLMMNIISK 19  
Db :||:| | | | | :||:|  
55 KRRIVDTNVLEGIQLIAK 73  
RESULT 25  
AAU78096  
ID AAU78096 standard; Protein; 76 AA.  
XX AAU78096;  
XX 18-JUN-2002 (first entry)  
XX Human DNA binding domain E2F-4.  
XX Human; telomerase reverse transcriptase; TERT; Site C; progeria; burn;  
XX repressor binding site; Hutchinson-Gilford syndrome; AIDS; cancer;  
XX acquired immunodeficiency syndrome; cardiovascular disease; osteoporosis;  
XX skin rejuvenation; immune senescence; bone marrow transplant; skin graft;  
XX neoplastic disease; TERT minimal promoter; DNA binding domain; E2F-4.  
XX Homo sapiens.  
XX OS OS  
XX Key Location/Qualifiers  
XX Domain 7..49  
XX /label= DNA\_binding\_domain  
XX WO200216657-A1.  
XX

```

PD XX 28-FEB-2002.
PF XX
XX PF 17-AUG-2001; 2001WO-US25861.
XX PF
PR XX 24-AUG-2000; 2000US-227865P.
PR XX 01-SEP-2000; 2000US-230174P.
PR XX 05-OCT-2000; 2000US-238345P.
XX XX
PA (SIER-) SIERRA SCI INC.
XX XX
PI Andrews WH, Foster CA, Fraser S, Mohammadpour H;
XX XX
PI WPI; 2002-280952/32.
XX XX
PT Modulating expression of telomerase reverse transcriptase (TERT) in a
PT cell, for regulating proliferative capacity of a cell, involves
PT modulating TERT transcription repression by Site C repressor binding
PT site -
XX XX
PS Disclosure; Page 8; 66pp; English.
XX XX
CC The present invention relates to a new method of modulating expression
CC of telomerase reverse transcriptase (TERT) from a TERT expression system
CC that includes a TERT promoter and a Site C repressor binding site. The
CC method of the invention involves modulating TERT transcription repression
CC by the Site C repressor binding site. The method of the invention is
CC useful for modulating expression of TERT for producing a mammalian
CC antibody. The method is also useful in a variety of different
CC applications, including immortalisation of cells, production of reagents
CC for use in life science research, therapeutic applications, and
CC therapeutic agent screening applications. Increasing TERT expression
CC delays natural telomeric shortening and/or increases telomeric length and
CC is useful for treating disease conditions such as Progeria or
CC Hutchinson-Gilford syndrome, acquired immunodeficiency syndrome (AIDS),
CC cardiovascular disease, osteoporosis, in skin rejuvenation and to inhibit
CC immune senescence. The method can be employed to lengthen telomeres of
CC osteoblast and osteoclast stem cells, encouraging bone replacement and
CC proper remodeling and reinforcement, and can thus be used in bone marrow
CC transplants for the treatment of cancer and skin grafts for burn
CC victims and as such the method improves the survival and effectiveness of
CC bone marrow and skin cell transplants. Decreasing TERT expression is
CC useful for treating cellular proliferative disease conditions, including
CC neoplastic disease conditions e.g. cancer. The present amino acid
CC sequence represents the human DNA binding domain E2F-4.
XX XX
SQ Sequence 76 AA;
Query Match 54.9%; Score 50; DB 23; Length 76;
Best Local Similarity 47.4%; Pred. No. 0.097;
Matches 9; Conservative 4; Mismatches 6; Indels 0; Gaps 0;

QY 1 RRRVYDALNVLMMNIISK 19
Db :||:||||| :||:
45 KRIYDITNVLEGIGLIEK 63

RESULT 26
ABB82987
ID ABB82987 standard; Protein; 76 AA.
XX AC ABB82987;
XX XX
XX 14-APR-2003 (first entry)
XX DE Repressor protein E2F-4 DNA binding domain.
XX XX
XX Telomerase reverse transcriptase; TERT; Site C repressor; transcription;
XX KW cytotostatic; immunostimulant; anti-HIV; vulnery; telomerase; human;
XX KW repressor protein; E2F-4.
XX XX
XX Homo sapiens.
XX OS
XX WO2002101010-A2.
XX PD

PD XX 19-DEC-2002.
XX XX
XX PF 06-JUN-2002; 2002WO-US17959.
XX PF
PR XX 07-JUN-2001; 2001US-296992P.
PR XX
XX (SIER-) SIERRA SCI INC.
XX XX
PI Foster CA, Fraser S, Mohammadpour H, Andrews WH;
XX XX
PI WPI; 2003-167401/16.
XX XX
PT Modulating expression of telomerase reverse transcriptase TERT by
PT blocking repression of TERT transcription, useful for the diagnosis and
PT treatment of disorders associated with aberrant telomerase activity
PT such as cancer and HIV -
XX XX
PS Disclosure; Page 7; 47pp; English.
XX XX
CC The invention relates to modulating expression of telomerase reverse
CC transcriptase (TERT) expression system that includes a TERT promoter and
CC a Site C repressor binding site. The method involves modulating TERT
CC transcription repression by the Site C repressor binding site. The
CC methods and compositions of the present invention are useful for the
CC immortalization of cells, production of reagents in life science
CC research, therapeutic agent screening applications, diagnosis and
CC treatment of disorders associated with aberrant telomerase activity such
CC as cancer, Progeria, immune senescence, HIV, and in skin rejuvenation.
CC The present sequence represents the DNA binding domain of a repressor
CC protein E2F-4.
XX XX
SQ Sequence 76 AA;
Query Match 54.9%; Score 50; DB 24; Length 76;
Best Local Similarity 47.4%; Pred. No. 0.097;
Matches 9; Conservative 4; Mismatches 6; Indels 0; Gaps 0;

QY 1 RRRVYDALNVLMMNIISK 19
Db :||:||||| :||:
45 KRIYDITNVLEGIGLIEK 63

RESULT 27
AAY44493
ID AAY44493 standard; peptide; 28 AA.
XX AC AAY44493;
XX XX
XX 27-MAR-2000 (first entry)
XX DT
XX DE Wheat E2F derived peptide-1.
XX XX
XX Wheat; transcription factor; TmE2F; plant retinoblastoma; plant E2F;
XX KW Rb protein; cell cycle regulation; plant growth regulation;
XX KW DNA replication; differentiation; E2F-derived peptide;
XX KW DNA binding domain.
XX XX
XX Triticum monococcum.
XX OS
XX Key Location/Qualifiers
XX FH Misc-difference 16 /label= Unknown
XX FT Misc-difference 21 /label= Unknown
XX FT Misc-difference 22 /label= Unknown
XX FT Misc-difference 25 /label= Unknown
XX FT
XX XX
XX WO9958681-A2.
XX PN
XX 18-NOV-1999.
XX PD

```

E2F activity inhibitors - for treatment and prevention of tumours and arteriosclerosis

Example 6; Page 31; 52pp; Japanese.

This represents a compound that can inhibit E2F activity. The compound is of the formula R1 - A - R2 where R1 is an optionally substituted alkanoyl, allyl, hetero-arylcarbonyl, alkoxy carbonyl, aryloxy carbonyl, hetero-aryloxy carbonyl, or H, R2 is OH, or optionally substituted alkoxy or amino, and A is an E2F family dimer forming region or DNA binding region, of at least 12 consecutive amino acids. Compounds of this formula can be used to inhibit E2F activity, and are useful in the treatment and prevention of tumours and arteriosclerosis.

Sequence 29 AA;

Query Match 50.5%; Score 46; DB 19; Length 29;  
Best Local Similarity 44.4%; Pred. NO. 0.16; 5; Indels 0; Gaps 0;  
Matches 8; Conservative 5; Mismatches

QY 1 RRRVDALNVLMAWNIIS 18  
          :|:|:| | | | : : :  
DB 12 KRRYDITNVLGIQLIA 29

RESULT 29  
AAW57049  
ID AAW57049 standard; peptide; 29 AA.  
XX AAW57049;  
AC  
XX 28-AUG-1998 (first entry)  
DT  
XX  
DE  
XX E2F activity inhibiting compound Ia-1.  
XX E2F activity; inhibitor; treatment; tumour; arteriosclerosis.  
XX Synthetic.  
OS Homo sapiens.  
XX  
Key Location/Qualifiers  
PH Modified-site 1  
FT /note= "N-terminal acetyl"  
FT Modified-site 29  
FT /note= "C-terminal amide"  
FT  
XX WO9814474-A1.  
PN  
XX  
PD 09-APR-1998.  
XX  
XX 26-SEP-1997; 97WO-JP03442.  
XX  
XX 30-SEP-1996; 96JP-0259432.  
XX  
XX (KYOW ) KYOWA HAKKO KOGYO KK.  
PA  
XX Mizukami T, Shibata K, Yamasaki M, Yoshida T;  
PI  
XX WPI; 1998-240020/21.  
DR  
XX  
XX E2F activity inhibitors - for treatment and prevention of tumours and arteriosclerosis  
PT  
XX  
XX Example 1; Page 24; 52pp; Japanese.

This represents a compound that can inhibit E2F activity. The compound is of the formula R1 - A - R2 where R1 is an optionally substituted alkanoyl, allyl, hetero-arylcarbonyl, alkoxy carbonyl, aryloxy carbonyl, hetero-aryloxy carbonyl or H, R2 is OH, or optionally substituted alkoxy or amino, and A is an E2F family dimer forming region or DNA binding region, of at least 12 consecutive amino acids. Compounds of this formula can be used to inhibit E2F activity, and are useful in the treatment and prevention of tumours and arteriosclerosis.

CC prevention of tumours and arteriosclerosis.

XX Sequence 29 AA;

SQ Query Match 50.5%; Score 46; DB 19; Length 29;

Best Local Similarity 44.4%; Pred. No. 0.16; Mismatches 5; Indels 0; Gaps 0;

Matches 8; Conservative

QY 1 RRVYDALNVLMAMNII 18

Db 12 KRIYDITNVLEGIQLIA 29

RESULT 30

AAU72601

ID AAU72601 standard; Peptide; 35 AA.

XX AAU72601;

AC AAU72601;

XX 26-FEB-2002 (first entry)

DT DEF domain consensus sequence.

XX Cell cycle protein; CCP; cell cycle regulation; herbicide;

KW plant growth regulator; CCP; cell cycle regulation; herbicide; biotic stress;

KW nutrient deprivation; pathogen attack; crop yield; motif.

XX Synthetic.

OS WO200185946-A2.

XX 15-NOV-2001.

PN 14-MAY-2001; 2001WO-IB01307.

XX 12-MAY-2000; 2000US-204045P.

XX (CROP-) CROPDESIGN NV.

XX Inze D, Boudolf V, De Veylder L, Acosta JAT, Magyar Z;

XX WPI; 2002-062249/08.

XX New cell cycle protein and nucleic acid molecule encoding it useful for

PT regulating cell cycle progression in plants and for identifying

PT modulators which are useful as herbicides or plant growth regulators -

XX Disclosure; Page 25; 316pp; English.

XX The invention relates to a novel cell cycle protein (CCP) and the

CC polynucleotides encoding them. CCP is useful for identifying a compound

CC which modulates the activity of the polypeptide and which binds to the

CC polypeptide and an anti-CCP antibody is useful for detecting the presence

CC of CCP in a sample. A CCP modulator is useful for modulating the cell

CC cycle or growth of a plant such as Arabidopsis thaliana, rice, wheat,

CC maize, tomato, alfalfa, oilseed rape, soybean, sunflower and canola.

CC CCP nucleic acid and polypeptide molecules are useful as modulating

CC agents in regulating cell cycle progression in plants. CCP is useful to

CC treat disorders characterised by insufficient or excessive production of

CC CCP protein or production of CCP protein forms which have decreased or

CC aberrant activity. Compounds that bind to or modulate the activity

CC of CCP polypeptide are useful as herbicides or plant growth regulators.

CC The polynucleotide is useful for modifying cell fate, plant development,

CC plant morphology, biochemistry and/or physiology, the length of the G1,

CC S, G2 and/or M phase of the cell cycle of a plant, initiation, promotion,

CC stimulation or enhancement of cell division, DNA replication, seed set,

CC seed size, seed development, tuber, fruit, leaf formation, shoot and root

CC initiation and/or development, nodule function, dwarfism in plants,

CC senescence, tolerance or resistance to stress. CCP, the polynucleotide

CC and the anti-CCP antibody are useful in agriculture to modulate the

CC protein levels or activity of a protein involved in the cell cycle due

CC to environmental conditions, including abiotic stress such as

CC cold, nutrient deprivation, heat, drought, salt stress, or biotic

CC stress such as pathogen attack, to modulate e.g. enhance crop yields,

CC and attenuate plant architecture, plant quality traits, plant

CC reproduction and seed development, endoreplication in storage cells,

CC storage tissues and/or storage organs of plants or its parts. CCP is

CC useful as an immunogen to generate antibodies. CCP protein is useful to

CC screen for naturally occurring CCP substrates. The polynucleotide is

CC useful for expressing CCP protein, to detect CCP mRNA, or a genetic

CC lesion in a CCP gene and to modulate CCP activity. The present sequence

CC represents a motif which may be found in a CCP protein of the invention.

XX

SQ Sequence 35 AA;

Query Match 49.5%; Score 45; DB 23; Length 35;

Best Local Similarity 64.7%; Pred. No. 0.32;

Matches 11; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 1 RRVYDALNVLMAMNII 17

Db 6 RXRXDNLNVXMAXXI 22

RESULT 31

AAW30502

ID AAW30502 standard; Peptide; 9 AA.

XX AAW30502;

AC AAW30502;

XX 26-OCT-1998 (first entry)

DT DP-1 transcription factor antagonist peptide motif.

XX DP-1; transcription factor; antagonist; E2F protein; apoptosis;

XX cell proliferation; cardiovascular cell; restenosis; tumour;

XX surgical stent; therapy.

XX Synthetic.

OS Homo sapiens.

XX WO9828334-A1.

XX 02-JUL-1998.

XX 22-DEC-1997; 97WO-GB03506.

XX 20-DEC-1996; 96GB-0026589.

XX (PROL-) PROLIFIX LTD.

XX Bandara LR, La Thangue NB;

XX WPI; 1998-377596/32.

XX Polypeptide fragments of the DP-1 transcription factor - used for

XX inducing apoptosis, specifically in tumour and cardiovascular cells,

XX e.g. for preventing restenosis

XX Claim 3; Page 44; 55pp; English.

XX This peptide comprises amino acid residues 175-183 in the DEF box

XX region (see AAW30501) of transcription factor DPL. Claimed peptides

XX (II) (see AAW30502-07) containing this and/or another motif (see

XX AAW30503) of the DEF box, are antagonists of the heterodimerisation

XX of a DP protein with an E2F protein. Also claimed are variants

XX of these peptides, especially containing substitutions of residues

XX corresponding to residues 167, 169, 171 and 175 of DP-1, fusion

XX proteins (III) comprising (I) or (II) and a membrane translocation

XX sequence (see AAW30508), expression vectors encoding (I)-(III) and

XX host cells. (I)-(III) are used therapeutically to induce apoptosis,

XX specifically in tumour or cardiovascular cells, either in vivo or in

XX vitro, e.g. for purging bone marrow. Surgical stents comprising

XX (I)-(III) are used to treat or prevent restenosis in patients who

XX have undergone angioplasty. (I)-(III) function by inactivating

XX the DNA-binding activity of DP/E2F heterodimers. They are also



GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: February 17, 2004, 10:50:13 ; Search time 13.7337 Seconds  
(without alignments)  
58.540 Million cell updates/sec

Title: US-09-900-147-3

Perfect score: 91

Sequence: 1 RRRVDALNVLMMNIISK 19

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 328717 seqs, 42310858 residues

Total number of hits satisfying chosen parameters: 228043

Minimum DB seq length: 0

Maximum DB seq length: 100

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued Patents AA:\*

- 1: /cgn2\_6/ptodata/1/iaa/5A COMB.pep.\*
- 2: /cgn2\_6/ptodata/1/iaa/5B COMB.pep.\*
- 3: /cgn2\_6/ptodata/1/iaa/6A COMB.pep.\*
- 4: /cgn2\_6/ptodata/1/iaa/6B COMB.pep.\*
- 5: /cgn2\_6/ptodata/1/iaa/PCTUS COMB.pep.\*
- 6: /cgn2\_6/ptodata/1/iaa/backfiles1.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	91	100.0	19	3	US-09-308-935-3
2	91	100.0	37	3	US-09-308-935-1
3	91	100.0	72	2	US-08-428-131-11
4	91	100.0	72	3	US-09-078-596-11
5	91	100.0	74	4	US-08-894-139-10
6	83	91.2	19	3	US-09-308-935-15
7	81	89.0	17	2	US-08-428-131-13
8	81	89.0	17	3	US-09-078-596-13
9	77	86.8	19	3	US-09-308-935-17
10	77	86.8	19	3	US-09-308-935-16
11	76	83.5	16	3	US-09-308-935-5
12	72	79.1	30	3	US-09-308-935-6
13	69	75.8	14	3	US-09-308-935-11
14	54	59.3	11	3	US-09-308-935-9
15	51	56.0	20	3	US-09-308-935-4
16	51	56.0	73	2	US-08-428-131-12
17	51	56.0	73	3	US-09-078-596-12
18	51	56.0	74	4	US-08-894-139-5
19	51	56.0	74	4	US-08-894-139-7
20	51	56.0	75	4	US-08-894-139-9
21	50	54.9	69	4	US-08-894-139-8
22	49	53.8	74	4	US-08-894-139-6
23	44	48.4	17	2	US-08-428-131-14
24	44	48.4	17	3	US-09-078-596-14
25	42	46.2	9	3	US-09-308-935-2
26	40	44.0	25	3	US-08-604-965E-1
27	40	44.0	40	3	US-08-604-965E-2

28	40	44.0	42	6	5258287-4	Patent No. 5258287
29	40	44.0	85	3	US-08-604-965E-9	Sequence 9, Appl1
30	37	40.7	60	6	5258287-1	Patent No. 5258287
31	36	39.6	54	4	US-09-205-258-455	Sequence 455, App
32	35	38.5	67	4	US-09-134-001C-4216	Sequence 4216, Ap
33	34	37.4	63	1	US-08-194-338-14	Sequence 14, Appl
34	34	37.4	88	4	US-08-936-165A-513	Sequence 513, App
35	32	35.2	25	1	US-08-378-761A-30	Sequence 30, Appl
36	32	35.2	25	1	US-08-485-286-30	Sequence 30, Appl
37	32	35.2	25	6	5248606-16	Patent No. 5248606
38	32	35.2	97	4	US-09-205-258-812	Sequence 812, App
39	31	34.1	7	3	US-09-308-935-7	Sequence 7, Appl1
40	31	34.1	22	3	US-08-604-965E-7	Sequence 7, Appl1
41	31	34.1	43	4	US-09-217-293-8	Sequence 8, Appl1
42	31	34.1	46	3	US-08-740-644-7	Sequence 7, Appl1
43	31	34.1	49	4	US-09-205-258-525	Sequence 525, App
44	31	34.1	56	4	US-09-732-210-903	Sequence 903, App
45	31	34.1	63	4	US-09-107-532A-5706	Sequence 5706, Ap

#### ALIGNMENTS

RESULT 1  
US-09-308-935-3  
; Sequence 3, Application US/09308935  
; Patent No. 6268334  
; GENERAL INFORMATION:  
; APPLICANT: La Thangue, Nicholas B  
; APPLICANT: Bandara, Lasantha R  
; TITLE OF INVENTION: Peptide antagonists of DP transcription factors  
; FILE REFERENCE: 620-67  
; CURRENT APPLICATION NUMBER: US/09/308,935  
; CURRENT FILING DATE: 1999-05-27  
; EARLIER APPLICATION NUMBER: PCT/GB97/03506  
; EARLIER FILING DATE: 1997-12-22  
; EARLIER APPLICATION NUMBER: GB 9626589.7  
; EARLIER FILING DATE: 1996-12-20  
; NUMBER OF SEQ ID NOS: 18  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 3  
; LENGTH: 19  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Description of Artificial Sequence: Synthetic  
US-09-308-935-3

Query Match 100.0%; Score 91; DB 3; Length 19;  
Best Local Similarity 100.0%; Pred. No. 7.3e-10;  
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 RRRVDALNVLMMNIISK 19  
Db 1 RRRVDALNVLMMNIISK 19

RESULT 2  
US-09-308-935-1  
; Sequence 1, Application US/09308935  
; Patent No. 6268334  
; GENERAL INFORMATION:  
; APPLICANT: La Thangue, Nicholas B  
; APPLICANT: Bandara, Lasantha R  
; TITLE OF INVENTION: Peptide antagonists of DP transcription factors  
; FILE REFERENCE: 620-67  
; CURRENT APPLICATION NUMBER: US/09/308,935  
; CURRENT FILING DATE: 1999-05-27  
; EARLIER APPLICATION NUMBER: PCT/GB97/03506  
; EARLIER FILING DATE: 1997-12-22  
; EARLIER APPLICATION NUMBER: GB 9626589.7  
; EARLIER FILING DATE: 1996-12-20  
; NUMBER OF SEQ ID NOS: 18

us-09-900-147-3.ra1

Tue Feb 17 11:55:48 2004

```

; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1
; LENGTH: 37
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Synthetic
US-09-308-935-1

Query Match      100.0%; Score 91; DB 3; Length 37;
Best Local Similarity 100.0%; Pred. No. 1.6e-09;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 RRRVYDALNVLMMNIISK 19
      |||||
DB      4 RRRVYDALNVLMMNIISK 22

RESULT 3
US-08-428-131-11
; Sequence 11, Application US/08428131
; Patent No. 5863757
; GENERAL INFORMATION:
; APPLICANT: La Thangue, Nicholas Barrie
; TITLE OF INVENTION: Transcription Factor DP-1
; NUMBER OF SEQUENCES: 14
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Nixon & Vanderhye
; STREET: 1100 No. 5863757th Glebe Road, 8th Floor
; CITY: Arlington
; STATE: Virginia
; COUNTRY: U.S.A.
; ZIP: 22201-4714
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25 (EPO)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/078,596
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/428,131
; FILING DATE: 23-JUN-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Arthur R. Crawford
; REGISTRATION NUMBER: 25,327
; REFERENCE/DOCKET NUMBER: 117-181
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703) 816-4000
; TELEFAX: (703) 816-4100
; INFORMATION FOR SEQ ID NO: 11:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 72 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-09-078-596-11

Query Match      100.0%; Score 91; DB 3; Length 72;
Best Local Similarity 100.0%; Pred. No. 3.6e-09;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 RRRVYDALNVLMMNIISK 19
      |||||
DB      7 RRRVYDALNVLMMNIISK 25

RESULT 5
US-08-894-139-10
; Sequence 10, Application US/08894139
; Patent No. 6448376
; GENERAL INFORMATION:
; APPLICANT: LA THANGUE, NICHOLAS B.
; APPLICANT: BERNARDS, RENE
; APPLICANT: HIJMAN, ELEANORE M.
; TITLE OF INVENTION: TRANSCRIPTION FACTOR E2F-5
; NUMBER OF SEQUENCES: 25
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: NIXON & VANDERHYE P.C.
; STREET: 1100 NORTH GLEBE ROAD
; CITY: ARLINGTON
; STATE: VIRGINIA
; COUNTRY: U.S.A.
; ZIP: 22201-4714
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/894,139
; FILING DATE: 13-AUG-1997
; CLASSIFICATION: 536
; ATTORNEY/AGENT INFORMATION:

```



NAME: WILSON, MARY J.  
REGISTRATION NUMBER: 32,955  
REFERENCE/DOCKET NUMBER: 620-22  
TELEPHONE: (703) 816-4000  
TELEFAX: (703) 816-4100  
INFORMATION FOR SEQ ID NO: 10:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 74 amino acids  
TYPE: amino acid  
STRANDEDNESS:  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
US-08-894-139-10

Query Match 100.0%; Score 91; DB 4; Length 74;  
Best Local Similarity 100.0%; Pred. No. 3.7e-09;  
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RRRVDALNVLMMNIISK 19  
Db 46 RRRVDALNVLMMNIISK 64

RESULT 6  
US-09-308-935-15  
Sequence 15, Application US/09308935  
Patent No. 6268334  
GENERAL INFORMATION:  
APPLICANT: La Thangue, Nicholas B  
APPLICANT: Bandara, Lasantha R  
TITLE OF INVENTION: Peptide antagonists of DP transcription factors  
FILE REFERENCE: 620-67  
CURRENT APPLICATION NUMBER: US/09/308,935  
CURRENT FILING DATE: 1999-05-27  
EARLIER APPLICATION NUMBER: PCT/GB97/03506  
EARLIER FILING DATE: 1997-12-22  
EARLIER APPLICATION NUMBER: GB 9626589.7  
EARLIER FILING DATE: 1996-12-20  
NUMBER OF SEQ ID NOS: 18  
SOFTWARE: Patentin Ver. 2.1  
SEQ ID NO 15  
TYPE: PRT  
LENGTH: 19  
ORGANISM: Artificial Sequence  
FEATURE:  
OTHER INFORMATION: Description of Artificial Sequence: Mutant peptide  
US-09-308-935-15

Query Match 91.2%; Score 83; DB 3; Length 19;  
Best Local Similarity 89.5%; Pred. No. 1.9e-08;  
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 RRRVDALNVLMMNIISK 19  
Db 1 RRRVDALNVLMMNIISK 19

RESULT 7  
US-08-428-131-13  
Sequence 13, Application US/08428131  
Patent No. 5863757  
GENERAL INFORMATION:  
APPLICANT: La Thangue, Nicholas Barrie  
TITLE OF INVENTION: Transcription Factor DP-1  
NUMBER OF SEQUENCES: 14  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Nixon & Vanderhye  
STREET: 1100 No. 5863757th Glebe Road, 8th Floor  
CITY: Arlington  
STATE: Virginia  
COUNTRY: U.S.A.  
ZIP: 22201-4714

NAME: WILSON, MARY J.  
REGISTRATION NUMBER: 32,955  
REFERENCE/DOCKET NUMBER: 620-22  
TELEPHONE: (703) 816-4000  
TELEFAX: (703) 816-4100  
INFORMATION FOR SEQ ID NO: 10:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 74 amino acids  
TYPE: amino acid  
STRANDEDNESS:  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
US-08-894-139-10

Query Match 89.0%; Score 81; DB 2; Length 17;  
Best Local Similarity 100.0%; Pred. No. 3.8e-08;  
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 RRRVDALNVLMMNIIS 18  
Db 1 RRRVDALNVLMMNIIS 17

RESULT 8  
US-09-078-596-13  
Sequence 13, Application US/09078596  
Patent No. 6150116  
GENERAL INFORMATION:  
APPLICANT: La Thangue, Nicholas Barrie  
TITLE OF INVENTION: Transcription Factor DP-1  
NUMBER OF SEQUENCES: 14  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Nixon & Vanderhye  
STREET: 1100 No. 6150116th Glebe Road, 8th Floor  
CITY: Arlington  
STATE: Virginia  
COUNTRY: U.S.A.  
ZIP: 22201-4714

COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.25 (EPO)  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/428,131  
FILING DATE: 23-JUN-1995  
CLASSIFICATION: 514  
ATTORNEY/AGENT INFORMATION:  
NAME: Arthur R. Crawford  
REGISTRATION NUMBER: 25,327  
REFERENCE/DOCKET NUMBER: 117-181  
TELEPHONE: (703) 816-4000  
TELEFAX: (703) 816-4100  
INFORMATION FOR SEQ ID NO: 13:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 17 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-428-131-13

COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.25 (EPO)  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/078,596  
FILING DATE:  
CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US/08/428,131  
FILING DATE: 23-JUN-1995  
ATTORNEY/AGENT INFORMATION:  
NAME: Arthur R. Crawford  
REGISTRATION NUMBER: 25,327  
REFERENCE/DOCKET NUMBER: 117-181  
TELEPHONE: (703) 816-4000  
TELEFAX: (703) 816-4100  
INFORMATION FOR SEQ ID NO: 13:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 17 amino acids  
TYPE: amino acid  
STRANDEDNESS: single

us-09-900-147-3.ra1

Tue Feb 17 11:55:48 2004

```
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-09-078-596-13

Query Match      89.0%; Score 81; DB 3; Length 17;
Best Local Similarity 100.0%; Pred. No. 3.8e-08;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 RRVYDALNVLMMNIISK 18
Db 1 RRVYDALNVLMMNIIS 17

RESULT 9
US-09-308-935-17
; Sequence 17, Application US/09308935
; Patent No. 6268334
; GENERAL INFORMATION:
; APPLICANT: La Thangue, Nicholas B
; APPLICANT: Bandara, Lasantha R
; TITLE OF INVENTION: Peptide antagonists of DP transcription factors
; FILE REFERENCE: 620-67
; CURRENT APPLICATION NUMBER: US/09/308,935
; CURRENT FILING DATE: 1999-05-27
; EARLIER APPLICATION NUMBER: PCT/GB97/03506
; EARLIER FILING DATE: 1997-12-22
; EARLIER APPLICATION NUMBER: GB 9626589.7
; EARLIER FILING DATE: 1996-12-20
; NUMBER OF SEQ ID NOS: 18
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 17
; LENGTH: 19
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Mutant peptide
US-09-308-935-17

Query Match      86.8%; Score 79; DB 3; Length 19;
Best Local Similarity 89.5%; Pred. No. 9.9e-08;
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 RRVYDALNVLMMNIISK 19
Db 1 RRVYDARNVMMNIISK 19

RESULT 10
US-09-308-935-16
; Sequence 16, Application US/09308935
; Patent No. 6268334
; GENERAL INFORMATION:
; APPLICANT: La Thangue, Nicholas B
; APPLICANT: Bandara, Lasantha R
; TITLE OF INVENTION: Peptide antagonists of DP transcription factors
; FILE REFERENCE: 620-67
; CURRENT APPLICATION NUMBER: US/09/308,935
; CURRENT FILING DATE: 1999-05-27
; EARLIER APPLICATION NUMBER: PCT/GB97/03506
; EARLIER FILING DATE: 1997-12-22
; EARLIER APPLICATION NUMBER: GB 9626589.7
; EARLIER FILING DATE: 1996-12-20
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 16
; LENGTH: 19
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Mutant peptide
US-09-308-935-16

Query Match      84.6%; Score 77; DB 3; Length 19;
Best Local Similarity 100.0%; Pred. No. 3e-06;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 5 YDALNVLMMNIISK 19
Db 1 RRVYDALNVLMMNIIS 16

RESULT 11
US-09-308-935-5
; Sequence 5, Application US/09308935
; Patent No. 6268334
; GENERAL INFORMATION:
; APPLICANT: La Thangue, Nicholas B
; APPLICANT: Bandara, Lasantha R
; TITLE OF INVENTION: Peptide antagonists of DP transcription factors
; FILE REFERENCE: 620-67
; CURRENT APPLICATION NUMBER: US/09/308,935
; CURRENT FILING DATE: 1999-05-27
; EARLIER APPLICATION NUMBER: PCT/GB97/03506
; EARLIER FILING DATE: 1997-12-22
; EARLIER APPLICATION NUMBER: GB 9626589.7
; EARLIER FILING DATE: 1996-12-20
; NUMBER OF SEQ ID NOS: 18
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 5
; LENGTH: 16
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Synthetic
US-09-308-935-5

Query Match      83.5%; Score 76; DB 3; Length 16;
Best Local Similarity 100.0%; Pred. No. 2.7e-07;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 RRVYDALNVLMMNIIS 18
Db 1 RRVYDALNVLMMNIIS 16

RESULT 12
US-09-308-935-6
; Sequence 6, Application US/09308935
; Patent No. 6268334
; GENERAL INFORMATION:
; APPLICANT: La Thangue, Nicholas B
; APPLICANT: Bandara, Lasantha R
; TITLE OF INVENTION: Peptide antagonists of DP transcription factors
; FILE REFERENCE: 620-67
; CURRENT APPLICATION NUMBER: US/09/308,935
; CURRENT FILING DATE: 1999-05-27
; EARLIER APPLICATION NUMBER: PCT/GB97/03506
; EARLIER FILING DATE: 1997-12-22
; EARLIER APPLICATION NUMBER: GB 9626589.7
; EARLIER FILING DATE: 1996-12-20
; NUMBER OF SEQ ID NOS: 18
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 6
; LENGTH: 30
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Synthetic
US-09-308-935-6

Query Match      79.1%; Score 72; DB 3; Length 30;
Best Local Similarity 100.0%; Pred. No. 3e-06;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 5 YDALNVLMMNIISK 19
Db 1 RRVYDALNVLMMNIIS 16
```

Dbb 1 YDALNVLMMNIIISK 15

## RESULT 13

US-09-308-935-11  
; Sequence 11, Application US/09308935  
; Patent No. 6268334  
; GENERAL INFORMATION:  
; APPLICANT: La Thangue, Nicholas B  
; APPLICANT: Bandara, Lasantha R  
; TITLE OF INVENTION: Peptide antagonists of DP transcription factors  
; FILE REFERENCE: 620-67  
; CURRENT APPLICATION NUMBER: US/09/308,935  
; CURRENT FILING DATE: 1999-05-27  
; EARLIER APPLICATION NUMBER: PCT/GB97/03506  
; EARLIER FILING DATE: 1997-12-22  
; EARLIER APPLICATION NUMBER: GB 9626589.7  
; EARLIER FILING DATE: 1996-12-20  
; NUMBER OF SEQ ID NOS: 18  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 11  
; LENGTH: 14  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Description of Artificial Sequence: Synthetic  
US-09-308-935-11

Query Match 75.8%; Score 69; DB 3; Length 14;  
Best Local Similarity 100.0%; Pred. No. 4.1e-06;  
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 RRVYDALNVLMMN 15

Db 1 RRVYDALNVLMMN 14

## RESULT 14

US-09-308-935-9  
; Sequence 9, Application US/09308935  
; Patent No. 6268334  
; GENERAL INFORMATION:  
; APPLICANT: La Thangue, Nicholas B  
; APPLICANT: Bandara, Lasantha R  
; TITLE OF INVENTION: Peptide antagonists of DP transcription factors  
; FILE REFERENCE: 620-67  
; CURRENT APPLICATION NUMBER: US/09/308,935  
; CURRENT FILING DATE: 1999-05-27  
; EARLIER APPLICATION NUMBER: PCT/GB97/03506  
; EARLIER FILING DATE: 1997-12-22  
; EARLIER APPLICATION NUMBER: GB 9626589.7  
; EARLIER FILING DATE: 1996-12-20  
; NUMBER OF SEQ ID NOS: 18  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 9  
; LENGTH: 11  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Description of Artificial Sequence: Synthetic  
US-09-308-935-9

Query Match 59.3%; Score 54; DB 3; Length 11;  
Best Local Similarity 100.0%; Pred. No. 0.0014;  
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RRVYDALNVL 11

Db 1 RRVYDALNVL 11

## RESULT 15

US-09-308-935-4

; Sequence 4, Application US/09308935  
; Patent No. 6268334  
; GENERAL INFORMATION:  
; APPLICANT: La Thangue, Nicholas B  
; APPLICANT: Bandara, Lasantha R  
; TITLE OF INVENTION: Peptide antagonists of DP transcription factors  
; FILE REFERENCE: 620-67  
; CURRENT APPLICATION NUMBER: US/09/308,935  
; CURRENT FILING DATE: 1999-05-27  
; EARLIER APPLICATION NUMBER: PCT/GB97/03506  
; EARLIER FILING DATE: 1997-12-22  
; EARLIER APPLICATION NUMBER: GB 9626589.7  
; EARLIER FILING DATE: 1996-12-20  
; NUMBER OF SEQ ID NOS: 18  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 4  
; LENGTH: 20  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Description of Artificial Sequence: Synthetic  
US-09-308-935-4

Query Match 56.0%; Score 51; DB 3; Length 20;  
Best Local Similarity 100.0%; Pred. No. 0.01;  
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 9 NVLMAMNIIISK 19

Db 1 NVLMAMNIIISK 11

## RESULT 16

US-08-428-131-12  
; Sequence 12, Application US/08428131  
; Patent No. 5863757  
; GENERAL INFORMATION:  
; APPLICANT: La Thangue, Nicholas Barrie  
; TITLE OF INVENTION: Transcription Factor DP-1  
; NUMBER OF SEQUENCES: 14  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Nixon & Vanderhye  
; STREET: 1100 No. 5863757th Glebe Road, 8th Floor  
; CITY: Arlington  
; STATE: Virginia  
; COUNTRY: U.S.A.  
; ZIP: 22201-4714  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.25 (BPO)  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/428,131  
; FILING DATE: 23-JUN-1995  
; CLASSIFICATION: 514  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Arthur R. Crawford  
; REGISTRATION NUMBER: 25,327  
; REFERENCE/DOCKET NUMBER: 117-181  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (703) 816-4100  
; TELEFAX: (703) 816-4100  
; INFORMATION FOR SEQ ID NO: 12:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 73 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
US-08-428-131-12

Query Match 56.0%; Score 51; DB 2; Length 73;

```
Best Local Similarity 47.4%; Pred. No. 0.047; Indels 5; Mismatches 5; Gaps 0;
Matches 9; Conservative 5;

OY 1 RRVYDALNVLMMNIISK 19
Db 7 KRRIYDITNVLEGIQIAK 25

RESULT 17
US-09-078-596-12
; Sequence 12, Application US/09078596
; Patent No. 6150116
; GENERAL INFORMATION:
; APPLICANT: La Thangue, Nicholas Barrie
; TITLE OF INVENTION: Transcription Factor DP-1
; NUMBER OF SEQUENCES: 14
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Nixon & Vanderhye
; STREET: 1100 No. 6150116th Glebe Road, 8th Floor
; CITY: Arlington
; STATE: Virginia
; COUNTRY: U.S.A.
; ZIP: 22201-4714
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25 (EPO)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/078,596
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/428,131
; FILING DATE: 23-JUN-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Arthur R. Crawford
; REGISTRATION NUMBER: 25,327
; REFERENCE/DOCKET NUMBER: 117-181
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703) 816-4000
; TELEFAX: (703) 816-4100
; INFORMATION FOR SEQ ID NO: 12:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 73 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-09-078-596-12

Query Match 56.0%; Score 51; DB 3; Length 73;
Best Local Similarity 47.4%; Pred. No. 0.047; Indels 5; Mismatches 5; Gaps 0;
Matches 9; Conservative 5;

OY 1 RRVYDALNVLMMNIISK 19
Db 7 KRRIYDITNVLEGIQIAK 25

RESULT 18
US-08-894-139-5
; Sequence 5, Application US/08894139
; Patent No. 6448376
; GENERAL INFORMATION:
; APPLICANT: LA THANGUE, NICHOLAS B.
; APPLICANT: BERNARDS, RENE
; APPLICANT: HIJMAN, ELEANORE M.
; TITLE OF INVENTION: TRANSCRIPTION FACTOR E2F-5
; NUMBER OF SEQUENCES: 25
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: NIXON & VANDERHYE P.C.
; STREET: 1100 NORTH GLEBE ROAD
; CITY: ARLINGTON
; STATE: VIRGINIA
; COUNTRY: U.S.A.
; ZIP: 22201-4714
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/894,139
; FILING DATE: 13-AUG-1997
; CLASSIFICATION: 536
; ATTORNEY/AGENT INFORMATION:
; NAME: WILSON, MARY J.
; REGISTRATION NUMBER: 32,955
; REFERENCE/DOCKET NUMBER: 620-22
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703) 816-4000
; TELEFAX: (703) 816-4100
; INFORMATION FOR SEQ ID NO: 7:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 74 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-08-894-139-5

Query Match 56.0%; Score 51; DB 4; Length 74;
Best Local Similarity 47.4%; Pred. No. 0.048; Indels 5; Mismatches 5; Gaps 0;
Matches 9; Conservative 5;

OY 1 RRVYDALNVLMMNIISK 19
Db 45 KRRIYDITNVLEGIQIAK 63

RESULT 19
US-08-894-139-7
; Sequence 7, Application US/08894139
; Patent No. 6448376
; GENERAL INFORMATION:
; APPLICANT: LA THANGUE, NICHOLAS B.
; APPLICANT: BERNARDS, RENE
; APPLICANT: HIJMAN, ELEANORE M.
; TITLE OF INVENTION: TRANSCRIPTION FACTOR E2F-5
; NUMBER OF SEQUENCES: 25
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: NIXON & VANDERHYE P.C.
; STREET: 1100 NORTH GLEBE ROAD
; CITY: ARLINGTON
; STATE: VIRGINIA
; COUNTRY: U.S.A.
; ZIP: 22201-4714
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/894,139
; FILING DATE: 13-AUG-1997
; CLASSIFICATION: 536
; ATTORNEY/AGENT INFORMATION:
; NAME: WILSON, MARY J.
; REGISTRATION NUMBER: 32,955
; REFERENCE/DOCKET NUMBER: 620-22
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703) 816-4000
; TELEFAX: (703) 816-4100
; INFORMATION FOR SEQ ID NO: 7:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 74 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-08-894-139-7
```

```
; LENGTH: 74 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-08-894-139-7

Query Match          56.0%; Score 51; DB 4; Length 74;
Best Local Similarity 47.4%; Pred. No. 0.048;
Matches 9; Conservative 5; Mismatches 5; Indels 0; Gaps 0;

Qy 1 RRRVDALNVLMAMNIISK 19
Db 45 KRRYDITNVLEGIDLIKK 63

RESULT 20
US-08-894-139-9
; Sequence 9, Application US/08894139
; Patent No. 6448376
; GENERAL INFORMATION:
; APPLICANT: LA THANGUE, NICHOLAS B.
; APPLICANT: BERNARDS, RENE
; APPLICANT: HIJMAN, ELEANORE M.
; TITLE OF INVENTION: TRANSCRIPTION FACTOR E2F-5
; NUMBER OF SEQUENCES: 25
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: NIXON & VANDERHYE P.C.
; STREET: 1100 NORTH GLEBE ROAD
; CITY: ARLINGTON
; STATE: VIRGINIA
; COUNTRY: U.S.A.
; ZIP: 22201-4714
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/894,139
; FILING DATE: 13-AUG-1997
; CLASSIFICATION: 536
; ATTORNEY/AGENT INFORMATION:
; NAME: WILSON, MARY J.
; REGISTRATION NUMBER: 32,955
; REFERENCE/DOCKET NUMBER: 620-22
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703) 816-4000
; TELEFAX: (703) 816-4100
; INFORMATION FOR SEQ ID NO: 8:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 69 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-08-894-139-8

Query Match          54.9%; Score 50; DB 4; Length 69;
Best Local Similarity 47.4%; Pred. No. 0.067;
Matches 9; Conservative 4; Mismatches 6; Indels 0; Gaps 0;

Qy 1 RRRVDALNVLMAMNIISK 19
Db 40 KRRYDITNVLEGIDLIKK 58

RESULT 22
US-08-894-139-6
; Sequence 6, Application US/08894139
; Patent No. 6448376
; GENERAL INFORMATION:
; APPLICANT: LA THANGUE, NICHOLAS B.
; APPLICANT: BERNARDS, RENE
; APPLICANT: HIJMAN, ELEANORE M.
; TITLE OF INVENTION: TRANSCRIPTION FACTOR E2F-5
; NUMBER OF SEQUENCES: 25
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: NIXON & VANDERHYE P.C.
; STREET: 1100 NORTH GLEBE ROAD
; CITY: ARLINGTON
; STATE: VIRGINIA
; COUNTRY: U.S.A.
; ZIP: 22201-4714
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/894,139
; FILING DATE: 13-AUG-1997
; CLASSIFICATION: 536
; ATTORNEY/AGENT INFORMATION:
; NAME: WILSON, MARY J.
; REGISTRATION NUMBER: 32,955
; REFERENCE/DOCKET NUMBER: 620-22
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703) 816-4000
; TELEFAX: (703) 816-4100
; INFORMATION FOR SEQ ID NO: 9:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 75 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-08-894-139-9

Query Match          56.0%; Score 51; DB 4; Length 75;
Best Local Similarity 47.4%; Pred. No. 0.049;
Matches 9; Conservative 5; Mismatches 5; Indels 0; Gaps 0;

Qy 1 RRRVDALNVLMAMNIISK 19
Db 46 KRRYDITNVLEGIDLIKK 64

RESULT 21
US-08-894-139-8
; Sequence 8, Application US/08894139
; Patent No. 6448376
; GENERAL INFORMATION:
; APPLICANT: LA THANGUE, NICHOLAS B.
```

us-09-900-147-3-rai

Tue Feb 17 11:55:48 2004

```

; REGISTRATION NUMBER: 32,955
; REFERENCE/DOCKET NUMBER: 620-22
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703) 816-4000
; TELEFAX: (703) 816-4100
; INFORMATION FOR SEQ ID NO: 6:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 74 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-08-894-139-6

Query Match 53.8%; Score 49; DB 4; Length 74;
Best Local Similarity 47.1%; Pred. No. 0.11;
Matches 9; Conservative 4; Mismatches 6; Indels 0; Gaps 0;

QY 1 RRVYDALNVLMMNIISK 19
Db 45 KRIYDITNVLEGIQLIA 63

RESULT 23
US-08-428-131-14
; Sequence 14, Application US/08428131
; Patent No. 5863757
; GENERAL INFORMATION:
; APPLICANT: La Thangue, Nicholas Barrie
; TITLE OF INVENTION: Transcription Factor DP-1
; NUMBER OF SEQUENCES: 14
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Nixon & Vanderhye
; STREET: 1100 No. 586375th Glebe Road, 8th Floor
; CITY: Arlington
; STATE: Virginia
; COUNTRY: U.S.A.
; ZIP: 22201-4714
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25 (EPO)
; CURRENT APPLICATION DATA:
; FILING DATE: 23-JUN-1995
; CLASSIFICATION: 514
; ATTORNEY/AGENT INFORMATION:
; NAME: Arthur R. Crawford
; REGISTRATION NUMBER: 25,327
; REFERENCE/DOCKET NUMBER: 117-181
; TELEPHONE: (703) 816-4000
; TELEFAX: (703) 816-4100
; INFORMATION FOR SEQ ID NO: 14:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 17 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-09-078-596-14

Query Match 48.4%; Score 44; DB 3; Length 17;
Best Local Similarity 47.1%; Pred. No. 0.14;
Matches 8; Conservative 4; Mismatches 5; Indels 0; Gaps 0;

QY 2 RRVYDALNVLMMNIISK 18
Db 1 RRIYDITNVLEGIQLIA 17

RESULT 25
US-09-308-935-2
; Sequence 2, Application US/09308935
; Patent No. 6268334
; GENERAL INFORMATION:
; APPLICANT: La Thangue, Nicholas B
; APPLICANT: Bandara, Lisantha R
; TITLE OF INVENTION: Peptide antagonists of DP transcription factors
; FILE REFERENCE: 620-67
; CURRENT APPLICATION NUMBER: US/09/308,935
; CURRENT FILING DATE: 1999-05-27
; EARLIER APPLICATION NUMBER: PCT/GB97/03506
; EARLIER FILING DATE: 1997-12-22
; EARLIER APPLICATION NUMBER: GB 9626589.7
; EARLIER FILING DATE: 1996-12-20
; NUMBER OF SEQ ID NOS: 18
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2
; LENGTH: 9
; TYPE: PPT
; ORGANISM: Artificial Sequence
; FEATURE:

```



APPLICATION NUMBER: US/07/171,623  
FILING DATE: 22-MAR-1988  
SEQ ID NO: 4;  
LENGTH: 42

Query Match 44.0%; Score 40; DB 6; Length 42;  
Best Local Similarity 44.4%; Pred. No. 2.2;  
Matches 8; Conservative 4; Mismatches 6; Indels 0; Gaps 0;

1 RRRVYDALNVLMMNIIS 18  
2 9 RREMEDTLNHLKFLNVL 26

RESULT 29  
US-08-604-965E-9  
Sequence 9, Application US/08604965E  
Patent No. 6046033  
GENERAL INFORMATION:  
APPLICANT: Goto, Masaaki  
APPLICANT: Tsuda, Eisuke  
APPLICANT: Yano, Kazuki  
APPLICANT: Kobayashi, Fumie  
APPLICANT: Yamaguchi, Kyoji  
APPLICANT: Washida, Naohiro  
APPLICANT: Satake, Toshiko  
APPLICANT: Morinaga, Tomonori  
APPLICANT: Ueda, Masatsugu  
APPLICANT: Higashio, Kanji  
TITLE OF INVENTION: Basic Osteoblast Growth FactorII (bOGF-  
NUMBER OF SEQUENCES: 20  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Burgess, Ryan and Wayne  
STREET: 370 Lexington Avenue, Suite 2105  
CITY: NEW YORK  
STATE: NEW YORK  
COUNTRY: UNITED STATES OF AMERICA  
ZIP: 10017  
COMPUTER READABLE FORM:  
MEDIUM TYPE: 5 1/4 inch diskette  
COMPUTER: PC'S LIMITED SYSTEM 200  
OPERATING SYSTEM: DOS  
SOFTWARE: WORDPERFECT 5.1  
CURRENT APPLICATION DATA:  
FILING DATE:  
CLASSIFICATION: 435  
PRIOR APPLICATION NUMBER: PCT/JP95/01270  
FILING DATE: June 27, 1994  
ATTORNEY/AGENT INFORMATION:  
NAME: Wayne, Milton J.  
REGISTRATION NUMBER: 17,906  
REFERENCE/DOCKET NUMBER: U-Wp-5212  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 212-683-8150  
TELEFAX: 212-532-4285  
TELEX: 423794  
INFORMATION FOR SEQ ID NO: 9:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 85  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: protein

US-08-604-965E-9  
Query Match 44.0%; Score 40; DB 3; Length 85;  
Best Local Similarity 44.4%; Pred. No. 5.1;  
Matches 8; Conservative 4; Mismatches 6; Indels 0; Gaps 0;

1 RRRVYDALNVLMMNIIS 18  
2 9 RREMEDTLNHLKFLNVL 26

RESULT 30  
5258287-1  
Patent No. 5258287  
APPLICANT: BAXTER, ROBERT C.; WOOD, WILLIAM I.  
TITLE OF INVENTION: DNA ENCODING AND METHODS OF PRODUCTION  
OF INSULIN-LIKE GROWTH FACTOR BINDING PROTEIN BP53  
NUMBER OF SEQUENCES: 58  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/07/171,623  
FILING DATE: 22-MAR-1988  
SEQ ID NO: 1;  
LENGTH: 60

5258287-1  
Query Match 40.7%; Score 37; DB 6; Length 60;  
Best Local Similarity 38.9%; Pred. No. 12;  
Matches 7; Conservative 5; Mismatches 6; Indels 0; Gaps 0;

1 RRRVYDALNVLMMNIIS 18  
2 9 RREMEDTLNHLKFLNVL 26

RESULT 31  
US-09-205-258-455  
Sequence 455, Application US/09205258  
Patent No. 6525174  
GENERAL INFORMATION:  
APPLICANT: Young et al.  
TITLE OF INVENTION: 207 Human Secreted Proteins  
FILE REFERENCE: P2007P1  
CURRENT APPLICATION NUMBER: US/09/205,258  
CURRENT FILING DATE: 1998-12-04  
EARLIER APPLICATION NUMBER: PCT/US98/11422  
EARLIER FILING DATE: 1998-06-04  
EARLIER APPLICATION NUMBER: 60/048,885  
EARLIER FILING DATE: 1997-06-06  
EARLIER APPLICATION NUMBER: 60/049,375  
EARLIER FILING DATE: 1997-06-06  
EARLIER APPLICATION NUMBER: 60/048,881  
EARLIER FILING DATE: 1997-06-06  
EARLIER APPLICATION NUMBER: 60/048,880  
EARLIER FILING DATE: 1997-06-06  
EARLIER APPLICATION NUMBER: 60/048,896  
EARLIER FILING DATE: 1997-06-06  
EARLIER APPLICATION NUMBER: 60/049,020  
EARLIER FILING DATE: 1997-06-06  
EARLIER APPLICATION NUMBER: 60/048,876  
EARLIER FILING DATE: 1997-06-06  
EARLIER APPLICATION NUMBER: 60/048,895  
EARLIER FILING DATE: 1997-06-06  
EARLIER APPLICATION NUMBER: 60/048,884  
EARLIER FILING DATE: 1997-06-06  
EARLIER APPLICATION NUMBER: 60/048,894  
EARLIER FILING DATE: 1997-06-06  
EARLIER APPLICATION NUMBER: 60/048,971  
EARLIER FILING DATE: 1997-06-06  
EARLIER APPLICATION NUMBER: 60/048,964  
EARLIER FILING DATE: 1997-06-06  
EARLIER APPLICATION NUMBER: 60/048,882  
EARLIER FILING DATE: 1997-06-06  
EARLIER APPLICATION NUMBER: 60/048,899  
EARLIER FILING DATE: 1997-06-06  
EARLIER APPLICATION NUMBER: 60/048,893  
EARLIER FILING DATE: 1997-06-06  
EARLIER APPLICATION NUMBER: 60/048,900  
EARLIER FILING DATE: 1997-06-06  
EARLIER APPLICATION NUMBER: 60/048,901



EARLIER FILING DATE: 1997-06-06  
EARLIER APPLICATION NUMBER: 60/048,892  
EARLIER FILING DATE: 1997-06-06  
EARLIER APPLICATION NUMBER: 60/048,915  
EARLIER FILING DATE: 1997-06-06  
EARLIER APPLICATION NUMBER: 60/049,019  
EARLIER FILING DATE: 1997-06-06  
EARLIER APPLICATION NUMBER: 60/048,970  
EARLIER FILING DATE: 1997-06-06  
EARLIER APPLICATION NUMBER: 60/048,972  
EARLIER FILING DATE: 1997-06-06  
EARLIER APPLICATION NUMBER: 60/048,916  
EARLIER FILING DATE: 1997-06-06  
EARLIER APPLICATION NUMBER: 60/049,373  
EARLIER FILING DATE: 1997-06-06  
EARLIER APPLICATION NUMBER: 60/048,875  
EARLIER FILING DATE: 1997-06-06  
EARLIER APPLICATION NUMBER: 60/049,374  
EARLIER FILING DATE: 1997-06-06  
EARLIER APPLICATION NUMBER: 60/048,917  
EARLIER FILING DATE: 1997-06-06  
EARLIER APPLICATION NUMBER: 60/048,949  
EARLIER FILING DATE: 1997-06-06  
EARLIER APPLICATION NUMBER: 60/048,974  
EARLIER FILING DATE: 1997-06-06  
EARLIER APPLICATION NUMBER: 60/048,883  
EARLIER FILING DATE: 1997-06-06  
EARLIER APPLICATION NUMBER: 60/048,897  
EARLIER FILING DATE: 1997-06-06  
EARLIER APPLICATION NUMBER: 60/048,898  
EARLIER FILING DATE: 1997-06-06  
EARLIER APPLICATION NUMBER: 60/048,962  
EARLIER FILING DATE: 1997-06-06  
EARLIER APPLICATION NUMBER: 60/048,963  
EARLIER FILING DATE: 1997-06-06  
EARLIER APPLICATION NUMBER: 60/048,877  
EARLIER FILING DATE: 1997-06-06  
EARLIER APPLICATION NUMBER: 60/048,878  
EARLIER FILING DATE: 1997-06-06  
EARLIER APPLICATION NUMBER: 60/070,923  
EARLIER FILING DATE: 1997-12-18  
EARLIER APPLICATION NUMBER: 60/092,921  
EARLIER FILING DATE: 1998-07-15  
EARLIER APPLICATION NUMBER: 60/094,657  
EARLIER FILING DATE: 1998-07-30  
NUMBER OF SEQ ID NOS: 1227  
SOFTWARE: PatentIn Ver. 2.0  
SEQ ID NO 455  
LENGTH: 54  
TYPE: PRT  
ORGANISM: Homo sapiens  
US-09-205-258-455

Query Match 39.6%; Score 36; DB 4; Length 54;  
Best Local Similarity 41.2%; Pred. No. 15;  
Matches 7; Conservative 5; Mismatches 5; Indels 0; Gaps 0;

QY 3 RYVDALNVLMMNIISK 19  
||| : : : : :  
DB 35 RYVPAMHFLCVHIYSK 51

RESULT 32  
US-09-134-001C-4216  
Sequence 4216, Application US/09134001C  
Patent No. 6380370  
GENERAL INFORMATION:  
APPLICANT: Lynn Doucette-Stamm et al  
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO STAPHYLOCOCCUS  
FILE REFERENCE: GTC-007  
CURRENT APPLICATION NUMBER: US/09/134,001C  
CURRENT FILING DATE: 1998-08-13

PRIOR APPLICATION NUMBER: US 60/064,964  
PRIOR FILING DATE: 1997-11-08  
PRIOR APPLICATION NUMBER: US 60/055,779  
PRIOR FILING DATE: 1997-08-14  
NUMBER OF SEQ ID NOS: 5674  
SEQ ID NO 4216  
LENGTH: 67  
TYPE: PRT  
ORGANISM: Staphylococcus epidermidis  
US-09-134-001C-4216

Query Match 38.5%; Score 35; DB 4; Length 67;  
Best Local Similarity 40.0%; Pred. No. 30;  
Matches 6; Conservative 6; Mismatches 3; Indels 0; Gaps 0;

QY 4 VYDALNVLMMNIIS 18  
||| : : : : :  
DB 11 LYSKLNILMSIILN 25

RESULT 33  
US-08-194-338-14  
Sequence 14, Application US/08194338  
Patent No. 5474898  
GENERAL INFORMATION:  
APPLICANT: Venter, John C.  
APPLICANT: Fraser, Claire M.  
APPLICANT: McComb, William R.  
TITLE OF INVENTION: OCTOPAMINE RECEPTOR  
NUMBER OF SEQUENCES: 16  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Knobbe, Martens, Olson and Bear  
STREET: 620 Newport Center Drive, Sixteenth Floor  
CITY: Newport Beach  
STATE: CA  
COUNTRY: USA  
ZIP: 92660  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/194,338  
FILING DATE: 08-FEB-1994  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/676,174  
FILING DATE: 28-MAR-1991  
ATTORNEY/AGENT INFORMATION:  
NAME: Israel, Ned A.  
REGISTRATION NUMBER: 29,655  
REFERENCE/DOCKET NUMBER: NIH101.001DV1  
TELEPHONE: (619) 235-8550  
TELEFAX: (619) 235-0176  
INFORMATION FOR SEQ ID NO: 14:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 63 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
HYPOTHETICAL: NO  
ANTI-SENSE: NO  
FRAGMENT TYPE: internal  
US-08-194-338-14

Query Match 37.4%; Score 34; DB 1; Length 63;  
Best Local Similarity 60.0%; Pred. No. 42;  
Matches 6; Conservative 4; Mismatches 0; Indels 0; Gaps 0;

QY 9 NVLMAMNIIS 18

Db 3 NVLVALNILA 12

||||:||||:

US-08-936-165A-513

Sequence 513, Application US/08936165A

Patent No. 6348582

GENERAL INFORMATION:

APPLICANT: Black, Michael

APPLICANT: Burnham, Martin

APPLICANT: Hodgson, John

APPLICANT: Knowles, David

APPLICANT: Lonetto, Michael

APPLICANT: Nicholas, Richard

APPLICANT: Pratt, Julie

APPLICANT: Reichard, Richard

APPLICANT: Rosenberg, Martin

APPLICANT: Ward, Judith

TITLE OF INVENTION: No. 6348582el Prokaryotic Polynucleotides, Polypeptides and Their Uses

NUMBER OF SEQUENCES: 534

CORRESPONDENCE ADDRESS:

ADDRESSEE: Smithkline Beecham Corporation

STREET: 709 Swedeland Road

CITY: King of Prussia

STATE: PA

COUNTRY: USA

ZIP: 19406-0939

COMPUTER READABLE FORM:

MEDIUM TYPE: Diskette

COMPUTER: IBM Compatible

OPERATING SYSTEM: DOS

SOFTWARE: FastSeq for Windows Version 2.0

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/936.165A

FILING DATE: 24-SEP-1997

CLASSIFICATION: 536

PRIOR APPLICATION DATA:

PRIOR APPLICATION NUMBER: 60/027,032

FILING DATE: 24-SEP-1996

ATTORNEY/AGENT INFORMATION:

NAME: Gimmi, Edward R

REGISTRATION NUMBER: 38,891

REFERENCE/DOCKET NUMBER: P50549

TELECOMMUNICATION INFORMATION:

TELEPHONE: 610-270-4478

TELEFAX: 610-270-5090

TELEX:

INFORMATION FOR SEQ ID NO: 513:

SEQUENCE CHARACTERISTICS:

LENGTH: 88 amino acids

TYPE: amino acid

STRANDEDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: Protein

US-08-936-165A-513

Query Match 37.4%; Score 34; DB 4; Length 88;

Best Local Similarity 56.2%; Pred. No. 62;

Matches 9; Conservative 3; Mismatches 2; Indels 2; Gaps 1;

QY 4 VYDAL--NVLMMNII 17

Db 48 MYDAKMDNVLVPINII 63

RESULT 35

US-08-378-761A-30

Sequence 30, Application US/08378761A

Patent No. 5635384

GENERAL INFORMATION:

APPLICANT: WALSH, TERENCE A

APPLICANT: HEY, TIMOTHY D

APPLICANT: MORGAN, ALICE ER

TITLE OF INVENTION: RIBOSOME-INACTIVATING PROTEINS, INACTIVE PRECURSOR FORMS THEREOF, A PROCESS FOR MAKING A METHOD OF

NUMBER OF SEQUENCES: 81

CORRESPONDENCE ADDRESS:

ADDRESSEE: ANDREA T. BORUCKI

STREET: 9330 ZIONSVILLE ROAD

CITY: INDIANAPOLIS

STATE: IN

COUNTRY: US

ZIP: 46268

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: PatentIn Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/485.286

FILING DATE: 26-JAN-1995

CLASSIFICATION: 435

ATTORNEY/AGENT INFORMATION:

NAME: BORUCKI, ANDREA T

REGISTRATION NUMBER: 33651

REFERENCE/DOCKET NUMBER: 38272B

TELECOMMUNICATION INFORMATION:

TELEPHONE: (317) 337-4846

INFORMATION FOR SEQ ID NO: 30:

SEQUENCE CHARACTERISTICS:

LENGTH: 25 amino acids

TYPE: amino acid

STRANDEDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: protein

US-08-378-761A-30

Query Match 35.2%; Score 32; DB 1; Length 25;

Best Local Similarity 55.6%; Pred. No. 31;

Matches 5; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 3 RVYDALNVL 11

Db 11 KYDSLNI 19

RESULT 36

US-08-485-286-30

Sequence 30, Application US/08485286

Patent No. 5646026

GENERAL INFORMATION:

APPLICANT: WALSH, TERENCE A

APPLICANT: HEY, TIMOTHY D

APPLICANT: MORGAN, ALICE ER

TITLE OF INVENTION: RIBOSOME-INACTIVATING PROTEINS, INACTIVE PRECURSOR FORMS THEREOF, A PROCESS FOR MAKING A METHOD OF

NUMBER OF SEQUENCES: 81

CORRESPONDENCE ADDRESS:

ADDRESSEE: ANDREA T. BORUCKI

STREET: 9330 ZIONSVILLE ROAD

CITY: INDIANAPOLIS

STATE: IN

COUNTRY: US

ZIP: 46268

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: PatentIn Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/485.286

FILING DATE: 26-JAN-1995

CLASSIFICATION: 435

ATTORNEY/AGENT INFORMATION:

NAME: BORUCKI, ANDREA T

REGISTRATION NUMBER: 33651

REFERENCE/DOCKET NUMBER: 38272B

TELECOMMUNICATION INFORMATION:

TELEPHONE: (317) 337-4846

INFORMATION FOR SEQ ID NO: 30:

SEQUENCE CHARACTERISTICS:

LENGTH: 25 amino acids

TYPE: amino acid

STRANDEDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: protein

US-08-378-761A-30

Query Match 35.2%; Score 32; DB 1; Length 25;

Best Local Similarity 55.6%; Pred. No. 31;

Matches 5; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 3 RVYDALNVL 11

Db 11 KYDSLNI 19

```
/ FILING DATE:
/ CLASSIFICATION: 435
/ PRIOR APPLICATION DATA:
/ APPLICATION NUMBER: US 08/378761
/ FILING DATE: 26-JAN-1995
/ ATTORNEY/AGENT INFORMATION:
/ NAME: BORUCKI, ANDREA T
/ REGISTRATION NUMBER: 33651
/ REFERENCE/DOCKET NUMBER: 38272B
/ TELECOMMUNICATION INFORMATION:
/ TELEPHONE: (317) 337-4846
/ INFORMATION FOR SEQ ID NO: 30:
/ SEQUENCE CHARACTERISTICS:
/ LENGTH: 25 amino acids
/ TYPE: amino acid
/ STRANDEDNESS: single
/ TOPOLOGY: linear
/ MOLECULE TYPE: protein
/ US-08-485-286-30

Query Match      35.2%; Score 32; DB 1; Length 25;
Best Local Similarity 55.6%; Pred. No. 31;
Matches 5; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 3 RYVDALNVL 11
Db 11 KTYDSLNI 19

RESULT 37
5248606-16
/ Patent No. 5248606
/ APPLICANT: WALSH, TERENCE A.;HEY, TIMOTHY D.;MORGAN,
/ ALICE E.R.
/ TITLE OF INVENTION: DNA ENCODING INACTIVE PRECURSOR AND
/ ACTIVE FORMS OF MAIZE RIBOSOME INACTIVATION
/ NUMBER OF SEQUENCES: 49
/ CURRENT APPLICATION DATA:
/ APPLICATION NUMBER: US/07/535,636
/ FILING DATE: 11-JUN-1990
/ SEQ ID NO:16:
/ LENGTH: 25
5248606-16

Query Match      35.2%; Score 32; DB 6; Length 25;
Best Local Similarity 55.6%; Pred. No. 31;
Matches 5; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 3 RYVDALNVL 11
Db 11 KTYDSLNI 19

RESULT 38
US-09-205-258-812
/ Sequence 812, Application US/09205258
/ Patent No. 6525174
/ GENERAL INFORMATION:
/ APPLICANT: Young et al.
/ TITLE OF INVENTION: 207 Human Secreted Proteins
/ FILE REFERENCE: P2007P1
/ CURRENT APPLICATION NUMBER: US/09/205,258
/ CURRENT FILING DATE: 1998-12-04
/ EARLIER APPLICATION NUMBER: PCT/US98/11422
/ EARLIER FILING DATE: 1998-06-04
/ EARLIER APPLICATION NUMBER: 60/048,885
/ EARLIER FILING DATE: 1997-06-06
/ EARLIER APPLICATION NUMBER: 60/049,375
/ EARLIER FILING DATE: 1997-06-06
/ EARLIER APPLICATION NUMBER: 60/048,881
/ EARLIER FILING DATE: 1997-06-06
/ EARLIER APPLICATION NUMBER: 60/048,880
/ EARLIER FILING DATE: 1997-06-06
/ EARLIER APPLICATION NUMBER: 60/048,896
/ EARLIER FILING DATE: 1997-06-06
/ EARLIER APPLICATION NUMBER: 60/049,020
/ EARLIER FILING DATE: 1997-06-06
/ EARLIER APPLICATION NUMBER: 60/048,876
/ EARLIER FILING DATE: 1997-06-06
/ EARLIER APPLICATION NUMBER: 60/048,895
/ EARLIER FILING DATE: 1997-06-06
/ EARLIER APPLICATION NUMBER: 60/048,884
/ EARLIER FILING DATE: 1997-06-06
/ EARLIER APPLICATION NUMBER: 60/048,894
/ EARLIER FILING DATE: 1997-06-06
/ EARLIER APPLICATION NUMBER: 60/048,971
/ EARLIER FILING DATE: 1997-06-06
/ EARLIER APPLICATION NUMBER: 60/048,964
/ EARLIER FILING DATE: 1997-06-06
/ EARLIER APPLICATION NUMBER: 60/048,882
/ EARLIER FILING DATE: 1997-06-06
/ EARLIER APPLICATION NUMBER: 60/048,899
/ EARLIER FILING DATE: 1997-06-06
/ EARLIER APPLICATION NUMBER: 60/048,893
/ EARLIER FILING DATE: 1997-06-06
/ EARLIER APPLICATION NUMBER: 60/048,900
/ EARLIER FILING DATE: 1997-06-06
/ EARLIER APPLICATION NUMBER: 60/048,901
/ EARLIER FILING DATE: 1997-06-06
/ EARLIER APPLICATION NUMBER: 60/048,892
/ EARLIER FILING DATE: 1997-06-06
/ EARLIER APPLICATION NUMBER: 60/048,915
/ EARLIER FILING DATE: 1997-06-06
/ EARLIER APPLICATION NUMBER: 60/049,019
/ EARLIER FILING DATE: 1997-06-06
/ EARLIER APPLICATION NUMBER: 60/048,970
/ EARLIER FILING DATE: 1997-06-06
/ EARLIER APPLICATION NUMBER: 60/048,972
/ EARLIER FILING DATE: 1997-06-06
/ EARLIER APPLICATION NUMBER: 60/048,916
/ EARLIER FILING DATE: 1997-06-06
/ EARLIER APPLICATION NUMBER: 60/049,373
/ EARLIER FILING DATE: 1997-06-06
/ EARLIER APPLICATION NUMBER: 60/048,875
/ EARLIER FILING DATE: 1997-06-06
/ EARLIER APPLICATION NUMBER: 60/049,374
/ EARLIER FILING DATE: 1997-06-06
/ EARLIER APPLICATION NUMBER: 60/048,917
/ EARLIER FILING DATE: 1997-06-06
/ EARLIER APPLICATION NUMBER: 60/048,949
/ EARLIER FILING DATE: 1997-06-06
/ EARLIER APPLICATION NUMBER: 60/048,974
/ EARLIER FILING DATE: 1997-06-06
/ EARLIER APPLICATION NUMBER: 60/048,883
/ EARLIER FILING DATE: 1997-06-06
/ EARLIER APPLICATION NUMBER: 60/048,897
/ EARLIER FILING DATE: 1997-06-06
/ EARLIER APPLICATION NUMBER: 60/048,898
/ EARLIER FILING DATE: 1997-06-06
/ EARLIER APPLICATION NUMBER: 60/048,962
/ EARLIER FILING DATE: 1997-06-06
/ EARLIER APPLICATION NUMBER: 60/048,963
/ EARLIER FILING DATE: 1997-06-06
/ EARLIER APPLICATION NUMBER: 60/048,877
/ EARLIER FILING DATE: 1997-06-06
/ EARLIER APPLICATION NUMBER: 60/048,878
/ EARLIER FILING DATE: 1997-06-06
/ EARLIER APPLICATION NUMBER: 60/070,923
/ EARLIER FILING DATE: 1997-12-18
/ EARLIER APPLICATION NUMBER: 60/092,921
/ EARLIER FILING DATE: 1998-07-15
/ EARLIER APPLICATION NUMBER: 60/094,657
/ NUMBER OF SEQ ID NOS: 1227
/ SOFTWARE: PatentIn Ver. 2.0
/ SEQ ID NO 812
```

us-09-900-147-3.ra1

Tue Feb 17 11:55:48 2004

```

; LENGTH: 97
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (16)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (38)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
US-09-205-258-812

Query Match          35.2%; Score 32; DB 4; Length 97;
Best Local Similarity 38.5%; Pred. No. 1.6e+02;
Matches 5; Conservative 5; Mismatches 3; Indels 0; Gaps 0;

QY 5 YDALNVLMMNII 17
Db 66 YDAIAVFLCIHIV 78

RESULT 39
US-09-308-935-7
; Sequence 7, Application US/09308935
; Patent No. 6268334
; GENERAL INFORMATION:
; APPLICANT: La Thangue, Nicholas B
; APPLICANT: Bandara, Lasantha R
; TITLE OF INVENTION: Peptide antagonists of DP transcription factors
; FILE REFERENCE: 620-67
; CURRENT APPLICATION NUMBER: US/09/308,935
; CURRENT FILING DATE: 1999-05-27
; EARLIER APPLICATION NUMBER: PCT/GB97/03506
; EARLIER FILING DATE: 1997-12-22
; EARLIER APPLICATION NUMBER: GB 9626589.7
; EARLIER FILING DATE: 1996-12-20
; NUMBER OF SEQ ID NOS: 18
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 7
; LENGTH: 7
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Synthetic
US-09-308-935-7

Query Match          34.1%; Score 31; DB 3; Length 7;
Best Local Similarity 100.0%; Pred. No. 2.5e+05;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 7 ALNVLMA 13
Db 1 ALNVLMA 7

RESULT 40
US-08-604-965E-7
; Sequence 7, Application US/08604965E
; Patent No. 6046033
; GENERAL INFORMATION:
; APPLICANT: Goto, Masaaki
; APPLICANT: Tenda, Eisuke
; APPLICANT: Yano, Kazuki
; APPLICANT: Kobayashi, Fumie
; APPLICANT: Yamaguchi, Kyoji
; APPLICANT: Washida, Naohiro
; APPLICANT: Satake, Toshiko
; APPLICANT: Morinaga, Tomonori
; APPLICANT: Ueda, Masatsugu
; APPLICANT: Higashio, Kanji
; TITLE OF INVENTION: Basic Osteoblast Growth FactorII (BOGF-
; II)

```

```

; NUMBER OF SEQUENCES: 20
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Burgess, Ryan and Wayne
; STREET: 370 Lexington Avenue, Suite 2105
; CITY: NEW YORK
; STATE: NEW YORK
; COUNTRY: UNITED STATES OF AMERICA
; ZIP: 10017
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 5 1/4 inch diskette
; COMPUTER: PC'S LIMITED SYSTEM 200
; OPERATING SYSTEM: DOS
; SOFTWARE: WORDPERFECT 5.1
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/604,965E
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/JP95/01270
; FILING DATE: June 27, 1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Wayne, Milton J.
; REGISTRATION NUMBER: 17,906
; REFERENCE/DOCKET NUMBER: U-Wp-5212
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 212-683-8150
; TELEFAX: 212-532-4285
; TELEX: 423794
; INFORMATION FOR SEQ ID NO: 7:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 22
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; US-08-604-965E-7

Query Match          34.1%; Score 31; DB 3; Length 22;
Best Local Similarity 46.2%; Pred. No. 40;
Matches 6; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

QY 6 DALNVLMMNIIIS 18
Db 1 DTLNHLKFLNVLIS 13

RESULT 41
US-09-217-293-8
; Sequence 8, Application US/09217293
; Patent No. 6337181
; GENERAL INFORMATION:
; APPLICANT: Stewart, Jeffrey J
; TITLE OF INVENTION: A METHOD OF SPECIFYING VACCINE COMPONENTS FOR VIRAL
; FILE REFERENCE: QUASISPECIES
; CURRENT APPLICATION NUMBER: US/09/217,293
; CURRENT FILING DATE: 1998-12-21
; NUMBER OF SEQ ID NOS: 9
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 8
; LENGTH: 43
; TYPE: PRT
; ORGANISM: Influenza A virus
; FEATURE:
; OTHER INFORMATION: Xaa1 is HY; Xaa2 is KE; Xaa3 is SRK; Xaa4 is DE;
; OTHER INFORMATION: Xaa5 is SN
; OTHER INFORMATION: Capital letters indicate advantageous variants and
; OTHER INFORMATION: lowercase letters indicate possibly advantageous
; OTHER INFORMATION: variants (see detailed description)
US-09-217-293-8

Query Match          34.1%; Score 31; DB 4; Length 43;
Best Local Similarity 63.6%; Pred. No. 90;
Matches 7; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

```

QY 5 YDALNVLMMN 15  
| | | | |  
Db 8 YPALNVTMPNN 18

## RESULT 42

US-08-740-644-7  
; Sequence 7, Application US/08740644  
; Patent No. 6033877  
; GENERAL INFORMATION:  
; APPLICANT: Timothy F. Murphy  
; APPLICANT: Kyungcheol Yi  
; TITLE OF INVENTION: Peptide Expression And Delivery System  
; NUMBER OF SEQUENCES: 8  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Hodgson, Russ, Andrews, Woods & Goodyear  
; STREET: 1800 One Mt Plaza  
; CITY: Buffalo  
; STATE: New York  
; COUNTRY: United States  
; ZIP: 14203-2391  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Diskette, 3.5 inch, 1.4 Mb storage  
; COMPUTER: IBM Compatible  
; OPERATING SYSTEM: MS-DOS/ Microsoft Windows 3.1  
; SOFTWARE: Wordperfect for Windows 5.1  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/740,644  
; FILING DATE: October 31, 1996  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 60/006,168  
; FILING DATE: No. 6033877ember 2, 1995  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Nelson, M. Bud  
; REGISTRATION NUMBER: 35,300  
; REFERENCE/DOCKET NUMBER: 11520.0065  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (716) 856-4000  
; TELEFAX: (716) 849-0349  
; INFORMATION FOR SEQ ID NO: 7:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 46 amino acid residues  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
; DESCRIPTION: protein  
; ORIGINAL SOURCE:  
; ORGANISM: H. influenzae  
US-08-740-644-7

Query Match 34.1%; Score 31; DB 3; Length 46;  
Best Local Similarity 50.0%; Pred.No. 98;  
Matches 5; Conservative 4; Mismatches 1; Indels 0; Gaps 0;

QY 5 YDALNVLMMN 14  
| | | | |  
Db 7 YDANNIIVAI 16

## RESULT 43

US-09-205-258-525  
; Sequence 525, Application US/09205258  
; Patent No. 6525174  
; GENERAL INFORMATION:  
; APPLICANT: Young et al.  
; TITLE OF INVENTION: 207 Human Secreted Proteins  
; FILE REFERENCE: P2007p1  
; CURRENT APPLICATION NUMBER: US/09/205,258  
; CURRENT FILING DATE: 1998-12-04  
; EARLIER APPLICATION NUMBER: PCT/US98/11422  
; EARLIER FILING DATE: 1998-06-04  
; EARLIER APPLICATION NUMBER: 60/048,885  
; EARLIER FILING DATE: 1997-06-06  
; EARLIER APPLICATION NUMBER: 60/049,375  
; EARLIER FILING DATE: 1997-06-06  
; EARLIER APPLICATION NUMBER: 60/048,881  
; EARLIER FILING DATE: 1997-06-06  
; EARLIER APPLICATION NUMBER: 60/048,880  
; EARLIER FILING DATE: 1997-06-06  
; EARLIER APPLICATION NUMBER: 60/048,896  
; EARLIER FILING DATE: 1997-06-06  
; EARLIER APPLICATION NUMBER: 60/049,020  
; EARLIER FILING DATE: 1997-06-06  
; EARLIER APPLICATION NUMBER: 60/048,876  
; EARLIER FILING DATE: 1997-06-06  
; EARLIER APPLICATION NUMBER: 60/048,895  
; EARLIER FILING DATE: 1997-06-06  
; EARLIER APPLICATION NUMBER: 60/048,884  
; EARLIER FILING DATE: 1997-06-06  
; EARLIER APPLICATION NUMBER: 60/048,894  
; EARLIER FILING DATE: 1997-06-06  
; EARLIER APPLICATION NUMBER: 60/048,971  
; EARLIER FILING DATE: 1997-06-06  
; EARLIER APPLICATION NUMBER: 60/048,964  
; EARLIER FILING DATE: 1997-06-06  
; EARLIER APPLICATION NUMBER: 60/048,882  
; EARLIER FILING DATE: 1997-06-06  
; EARLIER APPLICATION NUMBER: 60/048,899  
; EARLIER FILING DATE: 1997-06-06  
; EARLIER APPLICATION NUMBER: 60/048,893  
; EARLIER FILING DATE: 1997-06-06  
; EARLIER APPLICATION NUMBER: 60/048,900  
; EARLIER FILING DATE: 1997-06-06  
; EARLIER APPLICATION NUMBER: 60/048,901  
; EARLIER FILING DATE: 1997-06-06  
; EARLIER APPLICATION NUMBER: 60/048,892  
; EARLIER FILING DATE: 1997-06-06  
; EARLIER APPLICATION NUMBER: 60/048,915  
; EARLIER FILING DATE: 1997-06-06  
; EARLIER APPLICATION NUMBER: 60/049,019  
; EARLIER FILING DATE: 1997-06-06  
; EARLIER APPLICATION NUMBER: 60/048,970  
; EARLIER FILING DATE: 1997-06-06  
; EARLIER APPLICATION NUMBER: 60/048,972  
; EARLIER FILING DATE: 1997-06-06  
; EARLIER APPLICATION NUMBER: 60/048,916  
; EARLIER FILING DATE: 1997-06-06  
; EARLIER APPLICATION NUMBER: 60/049,373  
; EARLIER FILING DATE: 1997-06-06  
; EARLIER APPLICATION NUMBER: 60/048,875  
; EARLIER FILING DATE: 1997-06-06  
; EARLIER APPLICATION NUMBER: 60/049,374  
; EARLIER FILING DATE: 1997-06-06  
; EARLIER APPLICATION NUMBER: 60/048,917  
; EARLIER FILING DATE: 1997-06-06  
; EARLIER APPLICATION NUMBER: 60/048,949  
; EARLIER FILING DATE: 1997-06-06  
; EARLIER APPLICATION NUMBER: 60/048,974  
; EARLIER FILING DATE: 1997-06-06  
; EARLIER APPLICATION NUMBER: 60/048,883  
; EARLIER FILING DATE: 1997-06-06  
; EARLIER APPLICATION NUMBER: 60/048,897  
; EARLIER FILING DATE: 1997-06-06  
; EARLIER APPLICATION NUMBER: 60/048,898  
; EARLIER FILING DATE: 1997-06-06  
; EARLIER APPLICATION NUMBER: 60/048,962  
; EARLIER FILING DATE: 1997-06-06  
; EARLIER APPLICATION NUMBER: 60/048,963  
; EARLIER FILING DATE: 1997-06-06  
; EARLIER APPLICATION NUMBER: 60/048,877  
; EARLIER FILING DATE: 1997-06-06  
; EARLIER APPLICATION NUMBER: 60/048,878  
; EARLIER FILING DATE: 1997-06-06  
; EARLIER APPLICATION NUMBER: 60/070,923  
; EARLIER FILING DATE: 1997-12-18

```

; EARLIER APPLICATION NUMBER: 60/092,921
; EARLIER FILING DATE: 1998-07-15
; EARLIER APPLICATION NUMBER: 60/094,657
; EARLIER FILING DATE: 1998-07-30
; NUMBER OF SEQ ID NOS: 1227
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 525
; LENGTH: 49
; TYPE: PRT
; ORGANISM: Homo sapiens
; US-09-205-258-525

Query Match      34.1%; Score 31; DB 4; Length 49;
Best Local Similarity 37.5%; Pred. NO. 1.1e+02;
Matches 6; Conservative 2; Mismatches 8; Indels 0; Gaps 0;

QY      4 VYDALNVLMMNIISK 19
      :||:|:|
Db      34 IYDINMELMKRFSFK 49

RESULT 44
US-09-732-210-903
; Sequence 903, Application US/09732210
; Patent No. 6573361
; GENERAL INFORMATION:
; APPLICANT: Bunkers, Greg J.
; APPLICANT: Liang, Jihong
; APPLICANT: Mittanck, Cindy A.
; APPLICANT: Seale, Jeffrey W.
; APPLICANT: Wu, Yonnie S.
; TITLE OF INVENTION: Anti-fungal Proteins and Methods for Their Use
; FILE REFERENCE: 38-21(15036)B
; CURRENT APPLICATION NUMBER: US/09/732,210
; CURRENT FILING DATE: 2000-12-07
; PRIOR APPLICATION NUMBER: US 60/169,513
; PRIOR FILING DATE: 1999-12-07
; PRIOR APPLICATION NUMBER: US 60/169,340
; PRIOR FILING DATE: 1999-12-07
; NUMBER OF SEQ ID NOS: 1753
; SEQ ID NO 903
; LENGTH: 56
; TYPE: PRT
; ORGANISM: Escherichia coli
; US-09-732-210-903

Query Match      34.1%; Score 31; DB 4; Length 56;
Best Local Similarity 33.3%; Pred. NO. 1.2e+02;
Matches 5; Conservative 6; Mismatches 4; Indels 0; Gaps 0;

QY      2 RRVYDALNVLMMNI 16
      ||:|:|:|:|:|
Db      15 RRSHDALTAVTSLSV 29

RESULT 45
US-09-107-532A-5706
; Sequence 5706, Application US/09107532A
; Patent No. 6583275
; GENERAL INFORMATION:
; APPLICANT: Lynn A Doucette-Stamm and David Bush
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO
; ENTEROCOCCUS FAECIUM FOR DIAGNOSTICS AND THERAPEUTICS
; NUMBER OF SEQUENCES: 7310
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: GENOME THERAPEUTICS CORPORATION
; STREET: 100 Beaver Street
; CITY: Waltham
; STATE: Massachusetts
; COUNTRY: USA
; ZIP: 02354
; COMPUTER READABLE FORM:
; MEDIUM TYPE: CD/ROM ISO9660

```

```
Best Local Similarity 44.4%; Pred. No. 1.8;
Matches 8; Conservative 4; Mismatches 6; Indels 0; Gaps 0;

QY 1 RRRVYDALNVLMMNNIIS 18
   || : || | : ||
Db 8 RREMEDTLNHLKFLNVLIS 25

RESULT 34
AAR87181
ID AAR87181 standard; Peptide; 40 AA.
XX
AC AAR87181;
XX
DT 23-AUG-1996 (first entry)
XX
DE Osteoblast proliferating protein N-terminal fragment.
XX
KW Osteoblast proliferating protein; human; foetal fibroblast;
KW cell line IMR-90; ATCC-CCL186; treatment; bone reducing disease;
KW osteoporosis; antigen; immunological diagnosis; N-terminal.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT Misc-difference 30
FT /note= "undefined"
FT Misc-difference 35
FT /note= "undefined"
XX
PN WO9600240-A1.
XX
PD 04-JAN-1996.
XX
PF 26-JUN-1995; 95WO-JP01270.
XX
PR 27-JUN-1994; 94JP-0168984.
XX
XX (SNOW ) SNOW BRAND MILK PROD CO LTD.
XX
Goto M, Higashio K, Kobayashi F, Morinaga T, Satake T;
Tsuda E, Ueda M, Washida N, Yamaguchi K, Yano K;
WPI; 1996-068832/07.
XX
Novel 15 kD protein with osteoblast proliferating activity - useful
in treatment and diagnosis of bone reducing diseases such as
osteoporosis
Claim 4; Page 22; 41pp; Japanese.
XX
The present sequence is an N-terminal fragment from a protein which
has a mol. wt. of 15 kD by SDS-PAGE under reducing conditions,
strong affinity to cation exchanger and heparin and osteoblast
proliferating activity which is reduced after 10 mins. at 70
degrees C, and is lost after 10 mins. at 90 degrees C. The
protein, which was obtd. from the human foetal fibroblast cell
line IMR-90 (ATCC-CCL186), can be used to treat bone reducing
diseases, e.g. osteoporosis, and as an antigen for the
immunological diagnosis of such diseases.
XX
SQ Sequence 40 AA;
XX
Query Match 44.0%; Score 40; DB 17; Length 40;
Best Local Similarity 44.4%; Pred. No. 3.3;
Matches 8; Conservative 4; Mismatches 6; Indels 0; Gaps 0;

QY 1 RRRVYDALNVLMMNNIIS 18
   || : || | : ||
Db 8 RREMEDTLNHLKFLNVLIS 25

RESULT 35
AAR87182
ID AAR87182 standard; Protein; 85 AA.
XX
AC AAR87182;
XX
DT 23-AUG-1996 (first entry)
XX
DE Osteoblast proliferating protein.
XX
KW Osteoblast proliferating protein; human; foetal fibroblast;
KW cell line IMR-90; ATCC-CCL186; treatment; bone reducing disease;
KW osteoporosis; antigen; immunological diagnosis.
XX
OS Homo sapiens.
XX
PN WO9600240-A1.
XX
PD 04-JAN-1996.
XX
PF 26-JUN-1995; 95WO-JP01270.
XX
PR 27-JUN-1994; 94JP-0168984.
XX
XX (SNOW ) SNOW BRAND MILK PROD CO LTD.
XX
Goto M, Higashio K, Kobayashi F, Morinaga T, Satake T;
Tsuda E, Ueda M, Washida N, Yamaguchi K, Yano K;
WPI; 1996-068832/07.
XX
Novel 15 kD protein with osteoblast proliferating activity - useful
in treatment and diagnosis of bone reducing diseases such as
osteoporosis
Claim 5; Page 26; 41pp; Japanese.
XX
The present sequence is a protein, which has a mol. wt. of 15
kD by SDS-PAGE under reducing conditions, strong affinity to cation
exchanger and heparin and osteoblast proliferating activity which
is reduced after 10 mins. at 70 degrees C, and is lost after 10
mins. at 90 degrees C. The protein, which was obtd. from the
human foetal fibroblast cell line IMR-90 (ATCC-CCL186), can be used
to treat bone reducing diseases, e.g. osteoporosis, and as an
antigen for the immunological diagnosis of such diseases.
XX
SQ Sequence 85 AA;
XX
Query Match 44.0%; Score 40; DB 17; Length 85;
Best Local Similarity 44.4%; Pred. No. 8.7;
Matches 8; Conservative 4; Mismatches 6; Indels 0; Gaps 0;

QY 1 RRRVYDALNVLMMNNIIS 18
   || : || | : ||
Db 8 RREMEDTLNHLKFLNVLIS 25

RESULT 36
AAU07882
ID AAU07882 standard; Protein; 80 AA.
XX
AC AAU07882;
XX
DT 18-DEC-2001 (first entry)
XX
DE Polypeptide sequence for mammalian Sp987.
XX
KW Mammalian; reproductive-specific protein; male infertility;
KW spermatogenesis; sperm count disorder; anti infertility; reproduction.
XX
OS Mammalia.
XX
PN WO200166752-A2.
```

XX PD 13-SEP-2001.

XX XX 07-MAR-2001; 2001WO-US07371.

XX PF 07-MAR-2001; 2000US-0187518.

XX PR 12-JAN-2001; 2001US-0261557.

XX XX (WHED ) WHITEHEAD INST BIOMEDICAL RES.

XX PA Wang PJ, Page DC;

XX PI WPI; 2001-570774/64.

XX DR N-PSDB; AAS13646.

XX DR Novel reproduction-specific protein, useful for treating disorders of

XX PT reduced sperm count, enhancing/increasing sperm count and/or sperm

XX PT activity

XX XX Claim 22; Fig 48; 151pp; English.

XX XX The present invention relates to the isolation of novel mammalian and

XX CC human reproductive-specific proteins (AAU07859-AAU07899), and the

XX CC nucleic acids encoding them. The nucleic acids encoding

XX CC reproductive-specific proteins are useful for diagnosing infertility

XX CC which is a result of reduced sperm count, reduced sperm motility,

XX CC malformed sperm or combinations of these. The sequences of the invention

XX CC are useful as markers for spermatogonial cells, for identifying genes or

XX CC proteins characteristic of male infertility, diagnosing or aiding in

XX CC the diagnosis of infertility in men, and for contraception in which

XX CC sperm production or sperm count is reduced or defective sperm is

XX CC produced. Antibodies to reproductive-specific proteins are useful for

XX CC determining the presence of these proteins in a sample obtained from a

XX CC man being assessed for infertility, for identifying the expression of

XX CC genes in particular cell type or particular developmental stage, for

XX CC studies of spermatogenesis, and for immunofluorescence of germ cells or

XX CC binds. The sequences of the invention are also useful for treating

XX CC disorders of reduced sperm count, and for increasing sperm count and/or

XX CC sperm activity. The nucleic acids of the invention are useful in gene

XX CC therapy. AAU07859-AAU07882 represent the mammalian reproduction-specific

XX CC proteins of the present invention.

XX SQ Sequence 80 AA;

Query Match 42.9%; Score 39; DB 22; Length 80;

Best Local Similarity 53.3%; Pred. No. 12;

Matches 8; Conservative 1; Mismatches 6; Indels 0; Gaps 0;

QY 2 RRVYDALNVLMMNI 16

DB 30 RKYISVLNVAFACNI 44

RESULT 37

ABG66872

XX ID ABG66872 standard; Protein; 48 AA.

XX AC ABG66872;

XX XX 24-SRP-2002 (first entry)

XX DT Human prostate specific protein DEX0283\_180.

XX DE Human; prostate specific protein; cytostatic; PSP; prostate cancer;

XX KW gene therapy; vaccine; non-cancerous prostatic disease.

XX OS Homo sapiens.

XX XX WO200242329-A2.

XX PN 30-MAY-2002.

XX PD

XX XX

PF 20-NOV-2001; 2001WO-US45177.

XX PR 21-NOV-2000; 2000US-252189P.

XX PA (DIAD-) DIADEXUS INC.

XX XX Salceda S, Macina RA, Recipon H, Sun Y, Liu C;

XX PI WPI; 2002-500278/53.

XX DR New prostate specific polypeptides useful for identifying, diagnosing,

XX PT monitoring, staging, imaging and treating prostate cancer and

XX PT non-cancerous disease states in prostate

XX XX Claim 11; Page 233; 248pp; English.

XX XX The invention relates to an isolated polypeptide that comprises a

XX CC sequence with 60 % sequence identity to one of 103 prostate specific

XX CC polypeptide (PSP) sequences (S1), given in specification, or a sequence

XX CC encoded by a nucleic acid comprising one of 114 prostate specific nucleic

XX CC acid (PSNA) sequences (S2), given in specification. Also included are

XX CC a vector comprising the PSNA, a host cell comprising the vector,

XX CC preparation of the PSP, a polypeptide encoded by the PSNA; an anti-PSP

XX CC antibody the antibody or its fragment, a kit for detecting a risk of

XX CC cancer or presence of cancer in a patient, comprising a unit for

XX CC determining the presence of the PSP or the PSNA. The PSNA is useful for a patient

XX CC and a vaccine comprising the PSP or the PSNA. The PSNA is useful for

XX CC determining the presence of a prostate specific nucleic acid (PSNA) in a

XX CC sample. The antibody to the PSP is useful for determining the presence of

XX CC a prostate specific protein in a sample. The antibody is also useful for

XX CC treating a patient with prostate cancer, where the administration of the

XX CC antibody induces an immune response against the prostate cancer cell

XX CC expressing the nucleic acid molecule or polypeptide. The PSP or the PSNA

XX CC is useful for diagnosing and monitoring the presence and metastases of

XX CC prostate cancer in a patient. The PSP is useful as a vaccine component

XX CC for eliciting a humoral and/or cellular immune response, and for

XX CC identifying an agonist and antagonist. The PSNA is used in gene therapy,

XX CC producing transgenic animals and cells, and for producing engineered

XX CC prostate tissue for treatment and research. The PSNA is useful for

XX CC detecting prostate cancer by detecting genetic lesions or mutations. The

XX CC PSP and the PSNA are useful for detecting, diagnosing, monitoring,

XX CC staging and predicting cancers. The PSP and the PSNA are also useful for

XX CC identifying prostate tissue, and thus is useful in forensic science,

XX CC prostate cell differentiation and development and in tissue engineering.

XX CC The PSP, the PSNA and the antibody are useful for identifying,

XX CC diagnosing, monitoring, staging, imaging and treating prostate cancer and

XX CC non-cancerous disease states in prostate tissue. The present

XX CC sequence is a prostate specific protein (PSP) sequence.

XX SQ Sequence 48 AA;

Query Match 41.8%; Score 38; DB 23; Length 48;

Best Local Similarity 38.9%; Pred. No. 10;

Matches 7; Conservative 7; Mismatches 4; Indels 0; Gaps 0;

QY 2 RRVYDALNVLMMNIISK 19

DB 15 RRIYSALSLYISVLSK 32

RESULT 38

AAU22567

XX ID AAU22567 standard; Protein; 63 AA.

XX AC AAU22567;

XX XX 18-DEC-2001 (first entry)

XX DT Novel human colon associated polypeptide #100.

XX DE Human; colon cancer; congenital abnormality; infection; colitis;

XX KW inflammatory bowel disease; IBD; neoplastic disorder; gene therapy;

XX KW intestinal inflammatory disorder; malabsorption syndrome; gastric;

XX KW



KW sigmoid disease; antibacterial; antiviral; antiinflammatory;  
KW cytostatic.  
XX Homo sapiens.  
OS WO200155302-A2.  
XX PD 02-AUG-2001.  
XX PF 17-JAN-2001; 2001WO-US01240.  
XX PR 31-JAN-2000; 2000US-0179065.  
PR 04-FEB-2000; 2000US-0180828.  
PR 24-FEB-2000; 2000US-0184664.  
PR 02-MAR-2000; 2000US-0186350.  
PR 16-MAR-2000; 2000US-0189874.  
PR 17-MAR-2000; 2000US-0190076.  
PR 18-APR-2000; 2000US-0198123.  
PR 19-MAY-2000; 2000US-0205515.  
PR 07-JUN-2000; 2000US-0209467.  
PR 28-JUN-2000; 2000US-0214886.  
PR 30-JUN-2000; 2000US-0215135.  
PR 07-JUL-2000; 2000US-0216647.  
PR 07-JUL-2000; 2000US-0216880.  
PR 11-JUL-2000; 2000US-0217487.  
PR 11-JUL-2000; 2000US-0217496.  
PR 14-JUL-2000; 2000US-0218290.  
PR 26-JUL-2000; 2000US-0220963.  
PR 26-JUL-2000; 2000US-0220964.  
PR 14-AUG-2000; 2000US-0224518.  
PR 14-AUG-2000; 2000US-0224519.  
PR 14-AUG-2000; 2000US-0225213.  
PR 14-AUG-2000; 2000US-0225214.  
PR 14-AUG-2000; 2000US-0225266.  
PR 14-AUG-2000; 2000US-0225267.  
PR 14-AUG-2000; 2000US-0225268.  
PR 14-AUG-2000; 2000US-0225270.  
PR 14-AUG-2000; 2000US-0225447.  
PR 14-AUG-2000; 2000US-0225757.  
PR 14-AUG-2000; 2000US-0225758.  
PR 14-AUG-2000; 2000US-0225759.  
PR 18-AUG-2000; 2000US-0226279.  
PR 22-AUG-2000; 2000US-0226681.  
PR 22-AUG-2000; 2000US-0226868.  
PR 22-AUG-2000; 2000US-0227182.  
PR 23-AUG-2000; 2000US-0227009.  
PR 30-AUG-2000; 2000US-0228924.  
PR 01-SEP-2000; 2000US-0229287.  
PR 01-SEP-2000; 2000US-0229343.  
PR 01-SEP-2000; 2000US-0229344.  
PR 01-SEP-2000; 2000US-0229345.  
PR 05-SEP-2000; 2000US-0229509.  
PR 05-SEP-2000; 2000US-0229513.  
PR 06-SEP-2000; 2000US-0230437.  
PR 06-SEP-2000; 2000US-0230438.  
PR 08-SEP-2000; 2000US-0231242.  
PR 08-SEP-2000; 2000US-0231243.  
PR 08-SEP-2000; 2000US-0231244.  
PR 08-SEP-2000; 2000US-0231413.  
PR 08-SEP-2000; 2000US-0231414.  
PR 08-SEP-2000; 2000US-0232080.  
PR 08-SEP-2000; 2000US-0232081.  
PR 12-SEP-2000; 2000US-0231968.  
PR 14-SEP-2000; 2000US-0232397.  
PR 14-SEP-2000; 2000US-0232398.  
PR 14-SEP-2000; 2000US-0232399.  
PR 14-SEP-2000; 2000US-0232401.  
PR 14-SEP-2000; 2000US-0233063.  
PR 14-SEP-2000; 2000US-0233064.  
PR 14-SEP-2000; 2000US-0233065.  
PR 21-SEP-2000; 2000US-0234223.  
PR 21-SEP-2000; 2000US-0234274.

PR 25-SEP-2000; 2000US-0234997.  
PR 25-SEP-2000; 2000US-0234998.  
PR 26-SEP-2000; 2000US-0235484.  
PR 27-SEP-2000; 2000US-0235834.  
PR 27-SEP-2000; 2000US-0235836.  
PR 29-SEP-2000; 2000US-0236327.  
PR 29-SEP-2000; 2000US-0236367.  
PR 29-SEP-2000; 2000US-0236368.  
PR 29-SEP-2000; 2000US-0236369.  
PR 29-SEP-2000; 2000US-0236370.  
PR 02-OCT-2000; 2000US-0236802.  
PR 02-OCT-2000; 2000US-0237037.  
PR 02-OCT-2000; 2000US-0237038.  
PR 02-OCT-2000; 2000US-0237039.  
PR 02-OCT-2000; 2000US-0237040.  
PR 13-OCT-2000; 2000US-0239935.  
PR 13-OCT-2000; 2000US-0239937.  
PR 20-OCT-2000; 2000US-0240960.  
PR 20-OCT-2000; 2000US-0241221.  
PR 20-OCT-2000; 2000US-0241785.  
PR 20-OCT-2000; 2000US-0241786.  
PR 20-OCT-2000; 2000US-0241787.  
PR 20-OCT-2000; 2000US-0241808.  
PR 20-OCT-2000; 2000US-0241809.  
PR 20-OCT-2000; 2000US-0241826.  
PR 01-NOV-2000; 2000US-0244617.  
PR 08-NOV-2000; 2000US-0246474.  
PR 08-NOV-2000; 2000US-0246475.  
PR 08-NOV-2000; 2000US-0246476.  
PR 08-NOV-2000; 2000US-0246477.  
PR 08-NOV-2000; 2000US-0246478.  
PR 08-NOV-2000; 2000US-0246523.  
PR 08-NOV-2000; 2000US-0246524.  
PR 08-NOV-2000; 2000US-0246525.  
PR 08-NOV-2000; 2000US-0246526.  
PR 08-NOV-2000; 2000US-0246527.  
PR 08-NOV-2000; 2000US-0246528.  
PR 08-NOV-2000; 2000US-0246532.  
PR 08-NOV-2000; 2000US-0246609.  
PR 08-NOV-2000; 2000US-0246610.  
PR 08-NOV-2000; 2000US-0246611.  
PR 08-NOV-2000; 2000US-0246613.  
PR 17-NOV-2000; 2000US-0249207.  
PR 17-NOV-2000; 2000US-0249208.  
PR 17-NOV-2000; 2000US-0249209.  
PR 17-NOV-2000; 2000US-0249210.  
PR 17-NOV-2000; 2000US-0249211.  
PR 17-NOV-2000; 2000US-0249212.  
PR 17-NOV-2000; 2000US-0249213.  
PR 17-NOV-2000; 2000US-0249214.  
PR 17-NOV-2000; 2000US-0249215.  
PR 17-NOV-2000; 2000US-0249244.  
PR 17-NOV-2000; 2000US-0249245.  
PR 17-NOV-2000; 2000US-0249264.  
PR 17-NOV-2000; 2000US-0249265.  
PR 17-NOV-2000; 2000US-0249297.  
PR 17-NOV-2000; 2000US-0249299.  
PR 17-NOV-2000; 2000US-0249300.  
PR 01-DEC-2000; 2000US-0250160.  
PR 01-DEC-2000; 2000US-0250391.  
PR 05-DEC-2000; 2000US-0251030.  
PR 05-DEC-2000; 2000US-0251988.  
PR 05-DEC-2000; 2000US-0256719.  
PR 06-DEC-2000; 2000US-0251479.  
PR 06-DEC-2000; 2000US-0251856.  
PR 08-DEC-2000; 2000US-0251868.  
PR 08-DEC-2000; 2000US-0251869.  
PR 08-DEC-2000; 2000US-0251989.  
PR 08-DEC-2000; 2000US-0251990.  
PR 11-DEC-2000; 2000US-0254097.



PR 20-OCT-2000; 2000US-0241785.  
PR 20-OCT-2000; 2000US-0241786.  
PR 20-OCT-2000; 2000US-0241787.  
PR 20-OCT-2000; 2000US-0241808.  
PR 20-OCT-2000; 2000US-0241809.  
PR 20-OCT-2000; 2000US-0241826.  
PR 01-NOV-2000; 2000US-0244617.  
PR 08-NOV-2000; 2000US-0246474.  
PR 08-NOV-2000; 2000US-0246475.  
PR 08-NOV-2000; 2000US-0246476.  
PR 08-NOV-2000; 2000US-0246477.  
PR 08-NOV-2000; 2000US-0246478.  
PR 08-NOV-2000; 2000US-0246523.  
PR 08-NOV-2000; 2000US-0246524.  
PR 08-NOV-2000; 2000US-0246525.  
PR 08-NOV-2000; 2000US-0246526.  
PR 08-NOV-2000; 2000US-0246527.  
PR 08-NOV-2000; 2000US-0246528.  
PR 08-NOV-2000; 2000US-0246532.  
PR 08-NOV-2000; 2000US-0246609.  
PR 08-NOV-2000; 2000US-0246610.  
PR 08-NOV-2000; 2000US-0246611.  
PR 08-NOV-2000; 2000US-0246613.  
PR 17-NOV-2000; 2000US-0249207.  
PR 17-NOV-2000; 2000US-0249208.  
PR 17-NOV-2000; 2000US-0249209.  
PR 17-NOV-2000; 2000US-0249210.  
PR 17-NOV-2000; 2000US-0249211.  
PR 17-NOV-2000; 2000US-0249212.  
PR 17-NOV-2000; 2000US-0249213.  
PR 17-NOV-2000; 2000US-0249214.  
PR 17-NOV-2000; 2000US-0249215.  
PR 17-NOV-2000; 2000US-0249216.  
PR 17-NOV-2000; 2000US-0249217.  
PR 17-NOV-2000; 2000US-0249218.  
PR 17-NOV-2000; 2000US-0249244.  
PR 17-NOV-2000; 2000US-0249245.  
PR 17-NOV-2000; 2000US-0249264.  
PR 17-NOV-2000; 2000US-0249265.  
PR 17-NOV-2000; 2000US-0249297.  
PR 17-NOV-2000; 2000US-0249299.  
PR 17-NOV-2000; 2000US-0249300.  
PR 01-DEC-2000; 2000US-0250160.  
PR 01-DEC-2000; 2000US-0250391.  
PR 05-DEC-2000; 2000US-0251030.  
PR 05-DEC-2000; 2000US-0251988.  
PR 05-DEC-2000; 2000US-0256719.  
PR 06-DEC-2000; 2000US-0251479.  
PR 08-DEC-2000; 2000US-0251856.  
PR 08-DEC-2000; 2000US-0251868.  
PR 08-DEC-2000; 2000US-0251869.  
PR 08-DEC-2000; 2000US-0251989.  
PR 08-DEC-2000; 2000US-0251990.  
PR 11-DEC-2000; 2000US-0254097.  
PR 03-JAN-2001; 2001US-02559678.  
XX (HUMA-) HUMAN GENOME SCI INC.  
XX Rosen CA, Barash SC, Ruben SM;  
XX WPI; 2001-502630/55.  
XX N-PSDB; AAK88303.  
XX Polynucleotides encoding digestive system antigens, useful for  
PT diagnosing, treating, preventing and/or prognosing disorders of the  
PT digestive system, particularly cancer and cancer metastases -  
XX Claim 11; SEQ ID NO 1879; 986pp; English.  
XX The present invention provides the protein and coding sequences of a  
CC number of human digestive system antigens. These can be used in the  
CC diagnosis, treatment and prevention of digestive system disorders,  
CC including cancer, Meckel's diverticulum, bacterial or parasitic

CC infections, appendicitis, Hirschsprung's disease, chronic colitis or  
CC ulcerative colitis. The present sequence is a digestive system antigen of  
CC the invention.  
XX  
SQ Sequence 63 AA;  
Query Match 41.8%; Score 38; DB 22; Length 63;  
Best Local Similarity 56.2%; Pred. No. 14;  
Matches 9; Conservative 1; Mismatches 6; Indels 0; Gaps 0;  
QY 4 VYDALNVLMMNTISK 19  
Db 8 VYFLNFMKLNLFK 23  
RESULT 40  
AAR56646  
ID AAR56646 standard; peptide; 24 AA.  
XX  
AC AAR56646;  
XX  
25-MAR-2003 (updated)  
DT 23-MAR-1995 (first entry)  
XX  
Tyrosine activation motif mimic Fc-epsilon-RI-beta.  
XX  
Tyrosine activation motif mimic; arthritis; asthma; allergy;  
KW systemic lupus erythematosus; inflammatory bowe; disease;  
KW allograft rejection; antiinflammatory.  
XX  
Synthetic.  
OS  
WO9417095-A1.  
XX  
04-AUG-1994.  
XX  
28-JAN-1994; 94WO-US01025.  
XX  
29-JAN-1993; 93US-0013414.  
XX  
(ARIA-) ARIAD PHARM INC.  
XX  
Weigle M, Tao G, Sundaramoorthi R, Dalgarno DC, Zydowsky LD;  
PI Green J, Green OM;  
XX  
WPI; 1994-264027/32.  
XX  
New tyrosine activation motif analogues - for treating and  
PT preventing immune and inflammatory disorders, e.g. allergies, by  
PT inhibiting activation of mast cells, lymphocytes, etc.  
XX  
Claim 10; Page 75; 94pp; English.  
XX  
This TAM mimic may be used to treat or prevent conditions  
CC associated with immune responses and inflammation, e.g. arthritis,  
CC systemic lupus erythematosus, asthma, inflammatory bowel disease,  
CC allergies, allograft rejection, etc. This peptide probably prevents  
CC natural activation of target proteins so inhibit signal transduction.  
CC (Updated on 25-MAR-2003 to correct PN field.)  
CC (Updated on 25-MAR-2003 to correct PI field.)  
XX  
SQ Sequence 24 AA;  
Query Match 40.7%; Score 37; DB 15; Length 24;  
Best Local Similarity 63.8%; Pred. No. 6.3;  
Matches 7; Conservative 2; Mismatches 2; Indels 0; Gaps 0;  
QY 3 RYVDALNVLMA 13  
Db 3 RYVEELNLSA 13  
RESULT 41

Tue Feb 17 11:55:48 2004

us-09-900-147-3.rag

AAM89362  
ID AAM89362 standard; Protein; 87 AA.  
XX  
AC AAM89362;  
XX  
DT 07-NOV-2001 (first entry)  
XX  
DE Human immune/haematopoietic antigen SEQ ID NO:16955.  
XX  
KW Human; immune; haematopoietic; immune/haematopoietic antigen; cancer;  
KW cytostatic; gene therapy; vaccine; metastasis.  
XX  
OS Homo sapiens.  
XX  
FN WO200157182-A2.  
XX  
PD 09-AUG-2001.  
XX  
PF 17-JAN-2001; 2001WO-US01354.  
XX  
PR 31-JAN-2000; 2000US-0179065.  
PR 04-FEB-2000; 2000US-0180628.  
PR 24-FEB-2000; 2000US-0184664.  
PR 02-MAR-2000; 2000US-0186350.  
PR 16-MAR-2000; 2000US-0189874.  
PR 17-MAR-2000; 2000US-0190076.  
PR 18-APR-2000; 2000US-0198123.  
PR 19-MAY-2000; 2000US-0205515.  
PR 07-JUN-2000; 2000US-0209467.  
PR 28-JUN-2000; 2000US-0214986.  
PR 30-JUN-2000; 2000US-0215135.  
PR 07-JUL-2000; 2000US-0216647.  
PR 07-JUL-2000; 2000US-0216680.  
PR 11-JUL-2000; 2000US-0217487.  
PR 11-JUL-2000; 2000US-0217496.  
PR 14-JUL-2000; 2000US-0218290.  
PR 26-JUL-2000; 2000US-0220963.  
PR 14-AUG-2000; 2000US-0224518.  
PR 14-AUG-2000; 2000US-0224519.  
PR 14-AUG-2000; 2000US-0225213.  
PR 14-AUG-2000; 2000US-0225214.  
PR 14-AUG-2000; 2000US-0225266.  
PR 14-AUG-2000; 2000US-0225267.  
PR 14-AUG-2000; 2000US-0225268.  
PR 14-AUG-2000; 2000US-0225270.  
PR 14-AUG-2000; 2000US-0225447.  
PR 14-AUG-2000; 2000US-0225757.  
PR 14-AUG-2000; 2000US-0225758.  
PR 14-AUG-2000; 2000US-0225759.  
PR 18-AUG-2000; 2000US-0226279.  
PR 22-AUG-2000; 2000US-0226681.  
PR 22-AUG-2000; 2000US-0226868.  
PR 22-AUG-2000; 2000US-0227182.  
PR 23-AUG-2000; 2000US-0227009.  
PR 30-AUG-2000; 2000US-0228924.  
PR 01-SEP-2000; 2000US-0229287.  
PR 01-SEP-2000; 2000US-0229343.  
PR 01-SEP-2000; 2000US-0229344.  
PR 01-SEP-2000; 2000US-0229345.  
PR 05-SEP-2000; 2000US-0229509.  
PR 05-SEP-2000; 2000US-0229513.  
PR 06-SEP-2000; 2000US-0230437.  
PR 08-SEP-2000; 2000US-0230438.  
PR 08-SEP-2000; 2000US-0231242.  
PR 08-SEP-2000; 2000US-0231243.  
PR 08-SEP-2000; 2000US-0231244.  
PR 08-SEP-2000; 2000US-0231413.  
PR 08-SEP-2000; 2000US-0231414.  
PR 08-SEP-2000; 2000US-0232080.  
PR 08-SEP-2000; 2000US-0232081.  
PR 12-SEP-2000; 2000US-0231968.  
PR 14-SEP-2000; 2000US-0232397.  
PR 14-SEP-2000; 2000US-0232398.  
PR 14-SEP-2000; 2000US-0232399.  
PR 14-SEP-2000; 2000US-0232400.  
PR 14-SEP-2000; 2000US-0232401.  
PR 14-SEP-2000; 2000US-0233063.  
PR 14-SEP-2000; 2000US-0233064.  
PR 14-SEP-2000; 2000US-0233065.  
PR 21-SEP-2000; 2000US-0234223.  
PR 21-SEP-2000; 2000US-0234274.  
PR 25-SEP-2000; 2000US-0234997.  
PR 25-SEP-2000; 2000US-0234998.  
PR 26-SEP-2000; 2000US-0234984.  
PR 27-SEP-2000; 2000US-0235834.  
PR 27-SEP-2000; 2000US-0235836.  
PR 29-SEP-2000; 2000US-0236327.  
PR 29-SEP-2000; 2000US-0236367.  
PR 29-SEP-2000; 2000US-0236368.  
PR 29-SEP-2000; 2000US-0236369.  
PR 29-SEP-2000; 2000US-0236370.  
PR 02-OCT-2000; 2000US-0236802.  
PR 02-OCT-2000; 2000US-0237037.  
PR 02-OCT-2000; 2000US-0237038.  
PR 02-OCT-2000; 2000US-0237039.  
PR 13-OCT-2000; 2000US-0237040.  
PR 13-OCT-2000; 2000US-0239935.  
PR 20-OCT-2000; 2000US-0239937.  
PR 20-OCT-2000; 2000US-0240960.  
PR 20-OCT-2000; 2000US-0241221.  
PR 20-OCT-2000; 2000US-0241785.  
PR 20-OCT-2000; 2000US-0241786.  
PR 20-OCT-2000; 2000US-0241787.  
PR 20-OCT-2000; 2000US-0241808.  
PR 20-OCT-2000; 2000US-0241809.  
PR 20-OCT-2000; 2000US-0241826.  
PR 01-NOV-2000; 2000US-0244617.  
PR 08-NOV-2000; 2000US-0246474.  
PR 08-NOV-2000; 2000US-0246475.  
PR 08-NOV-2000; 2000US-0246476.  
PR 08-NOV-2000; 2000US-0246477.  
PR 08-NOV-2000; 2000US-0246478.  
PR 08-NOV-2000; 2000US-0246523.  
PR 08-NOV-2000; 2000US-0246524.  
PR 08-NOV-2000; 2000US-0246525.  
PR 08-NOV-2000; 2000US-0246526.  
PR 08-NOV-2000; 2000US-0246527.  
PR 08-NOV-2000; 2000US-0246528.  
PR 08-NOV-2000; 2000US-0246532.  
PR 08-NOV-2000; 2000US-0246609.  
PR 08-NOV-2000; 2000US-0246610.  
PR 08-NOV-2000; 2000US-0246611.  
PR 08-NOV-2000; 2000US-0246613.  
PR 17-NOV-2000; 2000US-0249207.  
PR 17-NOV-2000; 2000US-0249208.  
PR 17-NOV-2000; 2000US-0249209.  
PR 17-NOV-2000; 2000US-0249210.  
PR 17-NOV-2000; 2000US-0249211.  
PR 17-NOV-2000; 2000US-0249212.  
PR 17-NOV-2000; 2000US-0249213.  
PR 17-NOV-2000; 2000US-0249214.  
PR 17-NOV-2000; 2000US-0249215.  
PR 17-NOV-2000; 2000US-0249216.  
PR 17-NOV-2000; 2000US-0249217.  
PR 17-NOV-2000; 2000US-0249218.  
PR 17-NOV-2000; 2000US-0249244.  
PR 17-NOV-2000; 2000US-0249245.  
PR 17-NOV-2000; 2000US-0249264.  
PR 17-NOV-2000; 2000US-0249265.  
PR 17-NOV-2000; 2000US-0249297.  
PR 17-NOV-2000; 2000US-0249299.  
PR 17-NOV-2000; 2000US-0249300.  
PR 01-DEC-2000; 2000US-0250160.  
PR 01-DEC-2000; 2000US-0250391.  
PR 05-DEC-2000; 2000US-0251030.

PR 05-DEC-2000; 2000US-0251988.  
 PR 05-DEC-2000; 2000US-0256719.  
 PR 06-DEC-2000; 2000US-0251479.  
 PR 08-DEC-2000; 2000US-0251856.  
 PR 08-DEC-2000; 2000US-0251868.  
 PR 08-DEC-2000; 2000US-0251869.  
 PR 08-DEC-2000; 2000US-0251989.  
 PR 08-DEC-2000; 2000US-0251990.  
 PR 11-DEC-2000; 2000US-0254097.  
 PR 05-JAN-2001; 2001US-0259678.  
 PA (HUMA-) HUMAN GENOME SCI INC.  
 XX  
 XX Rosen CA, Barash SC, Ruben SM;  
 XX  
 XX WPI; 2001-483426/52.  
 DR N-PSDB; AAK62143.  
 XX  
 PT Nucleic acids encoding human immune/hematopoietic antigen polypeptides,  
 PT useful for preventing, diagnosing and/or treating cancers and  
 PT metastasis -  
 XX  
 PS Claim 11; SEQ ID NO 16955; 3071pp + Sequence Listing; English.  
 XX  
 CC AAK54951 to AAK64702 encode the human immune/hematopoietic antigen (I)  
 CC amino acid sequences given in AAK82170 to AAK91921. (I) have cytostatic  
 CC activity, and can be used in gene therapy and vaccine production. (I)  
 CC proteins and polynucleotides may be used in the prevention, diagnosis and  
 CC treatment of diseases associated with inappropriate (I) expression. For  
 CC example, they may be used to treat disorders associated with decreased  
 CC expression by rectifying mutations or deletions in a patient's genome  
 CC that affect the activity of (I) by expressing inactive proteins or to  
 CC supplement the patient's own production of (I). Additionally, (I)  
 CC polynucleotides may be used to produce the secreted (I), by inserting  
 CC the nucleic acids into a host cell and culturing the cell to express the  
 CC protein. (I) proteins and polynucleotides may be used to prevent,  
 CC diagnose and treat immune/hematopoietic-related diseases, especially  
 CC cancers and cancer metastases of haematopoietic-derived cells. AAK64703  
 CC to AAK87694 represent human immune/hematopoietic antigen genomic  
 CC sequences from the present invention. AAK54942 to AAK54950 and AAK82169  
 CC represent sequences used in the exemplification of the present invention.  
 XX  
 SQ Sequence 87 AA;  
 Query Match 40.7%; Score 37; DB 22; Length 87;  
 Best Local Similarity 53.3%; Pred. No. 33;  
 Matches 8; Conservative 1; Mismatches 6; Indels 0; Gaps 0;  
 QY 5 YDALNVLMMNNIISK 19  
 |||||  
 Db 73 YDILNFLNAKNTLK 87  
 RESULT 42  
 AAW57050  
 ID AAW57050 standard; peptide; 15 AA.  
 XX  
 AC AAW57050;  
 XX  
 DT 28-AUG-1998 (first entry)  
 XX  
 DE E2F activity inhibiting compound Ia-2.  
 XX  
 KW E2F activity; inhibitor; treatment; tumour; arteriosclerosis.  
 XX  
 OS Synthetic.  
 OS Homo sapiens.  
 XX  
 Key Location/Qualifiers  
 FH Modified-site 1  
 FT /note= "N-terminal acetyl"  
 FT 15  
 FT Modified-site  
 FT /note= "C-terminal amide"

XX WO9814474-A1.  
 XX  
 PD 09-APR-1998.  
 XX  
 PF 26-SEP-1997; 97WO-JP03442.  
 XX  
 PR 30-SEP-1996; 96JP-0259432.  
 XX  
 PA (KYOW ) KYOWA HAKKO KOGYO KK.  
 XX  
 PI Mizukami T, Shibata K, Yamasaki M, Yoshida T;  
 XX  
 DR WPI; 1998-240020/21.  
 XX  
 PT E2F activity inhibitors - for treatment and prevention of tumours  
 PT and arteriosclerosis  
 XX  
 PS Example 2; Page 26; 52pp; Japanese.  
 XX  
 CC This represents a compound that can inhibit E2F activity. The compound  
 CC is of the formula R1 - A - R2 where R1 is an optionally substituted  
 CC alkanoyl, allyl, hetero-arylcarbonyl, alkoxycarbonyl, aryloxycarbonyl,  
 CC hetero-aryloxycarbonyl, or H, R2 is OH, or optionally substituted alkoxy  
 CC or amino, and A is an E2F family dimer forming region or DNA binding  
 CC region, of at least 12 consecutive amino acids. Compounds of this formula  
 CC can be used to inhibit E2F activity, and are useful in the treatment and  
 CC prevention of tumours and arteriosclerosis.  
 XX  
 SQ Sequence 15 AA;  
 Query Match 39.6%; Score 36; DB 19; Length 15;  
 Best Local Similarity 60.0%; Pred. No. 5.4;  
 Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;  
 QY 1 RRRVYDALNV 10  
 :||:|  
 Db 6 KRIYDITNV 15  
 RESULT 43  
 ABG56095  
 ID ABG56095 standard; Peptide; 41 AA.  
 XX  
 AC ABG56095;  
 XX  
 DT 25-FEB-2003 (first entry)  
 XX  
 DE Human liver peptide, SEQ ID No 34743.  
 XX  
 KW Human; liver; cirrhosis; hyperlipoproteinaemia; hyperlipidaemia;  
 KW hypercholesterolaemia; coronary heart disease.  
 XX  
 OS Homo sapiens.  
 XX  
 PN WO200157273-A2.  
 XX  
 PD 09-AUG-2001.  
 XX  
 PF 30-JAN-2001; 2001WO-US00664.  
 XX  
 PR 04-FEB-2000; 2000US-0180312.  
 PR 26-MAY-2000; 2000US-0207456.  
 PR 30-JUN-2000; 2000US-0608408.  
 PR 03-AUG-2000; 2000US-0632366.  
 PR 21-SEP-2000; 2000US-0234687.  
 PR 27-SEP-2000; 2000US-0236359.  
 PR 04-OCT-2000; 2000US-0024263.  
 XX  
 PA (MOLE-) MOLECULAR DYNAMICS INC.  
 XX  
 PI Penn SG, Hanzel DK, Chen W, Rank DR;  
 XX

Tue Feb 17 11:55:48 2004

DR WPI; 2001-488898/53.  
 XX Human genome-derived single exon nucleic acid probes useful for  
 PT analysing gene expression in human adult liver -  
 XX  
 XX  
 PS Claim 27; SEQ ID No 34743; 658pp; English.  
 XX  
 CC The invention relates to a single exon nucleic acid probe (SENP) (I) for  
 CC measuring human gene expression in a sample derived from human adult  
 CC liver, comprising one of 13109 defined nucleotide sequences given in the  
 CC specification (or complements/ fragments). The probe hybridises at high  
 CC stringency to a nucleic acid molecule expressed in the human adult  
 CC liver. (I) may be used for predicting, measuring and displaying gene  
 CC expression in samples derived from human adult liver. The genes  
 CC identified may be involved in genetic liver diseases such as cirrhosis,  
 CC hyperlipoproteinaemia, hyperlipidaemia and hypercholesterolaemia which  
 CC is associated with coronary heart disease. ABG47348-ABG59930 represent  
 CC human liver single exon encoded peptides of the invention.  
 CC Note: The sequence information for this patent does not appear in the  
 CC printed specification but was obtained in electronic format directly  
 CC from WIPO at ftp.wipo.int/pub/published\_pct\_sequences.  
 XX  
 XX  
 SQ Sequence 41 AA;  
 Query Match 39.6%; Score 36; DB 22; Length 41;  
 Best Local Similarity 53.3%; Pred. No. 19;  
 Matches 8; Conservative 2; Mismatches 5; Indels 0; Gaps 0;  
 QY 2 RRVYDALNVLMMNI 16  
 DB 22 REYQDLLNVKWLDI 36  
 RESULT 44  
 ABB40647  
 ID ABB40647 standard; Peptide; 41 AA.  
 XX  
 AC ABB40647;  
 XX  
 XX  
 DT 04-FEB-2002 (first entry)  
 DE Peptide #8153 encoded by human foetal liver single exon probe.  
 XX Human; foetal liver; gene expression; single exon nucleic acid probe.  
 KW Homo sapiens.  
 OS  
 XX WO200157277-A2.  
 PN  
 PD 09-AUG-2001.  
 XX  
 PF 30-JAN-2001; 2001WO-US00669.  
 XX  
 XX 04-FEB-2000; 2000US-0180312.  
 PR 26-MAY-2000; 2000US-0207456.  
 PR 30-JUN-2000; 2000US-0608408.  
 PR 03-AUG-2000; 2000US-0632366.  
 PR 21-SEP-2000; 2000US-0234687.  
 PR 27-SEP-2000; 2000US-0236359.  
 PR 04-OCT-2000; 2000GB-0024263.  
 XX  
 XX (MOLE-) MOLECULAR DYNAMICS INC.  
 PA  
 XX Penn SG, Hanzel DK, Chen W, Rank DR;  
 PI  
 XX WPI; 2001-483447/52.  
 XX  
 DR Human genome-derived single exon nucleic acid probes useful for  
 XX analysing gene expression in human fetal liver -  
 PT  
 XX  
 PS Claim 27; SEQ ID NO 33282; 639pp + sequence listing; English.  
 XX  
 XX The invention relates to a single exon nucleic acid probe for

CC measuring human gene expression in a sample derived from human foetal  
 CC liver. The single exon nucleic acid probes may be used for predicting,  
 CC measuring and displaying gene expression in samples derived from human  
 CC fetal liver. The present sequence is a peptide encoded by a single exon  
 CC nucleic acid probe of the invention.  
 CC Note: The sequence data for this patent did not form part of the  
 CC printed specification, but was obtained in electronic format directly  
 CC from WIPO at ftp.wipo.int/pub/published\_pct\_sequences.  
 XX  
 XX  
 SQ Sequence 41 AA;  
 Query Match 39.6%; Score 36; DB 22; Length 41;  
 Best Local Similarity 53.3%; Pred. No. 19;  
 Matches 8; Conservative 2; Mismatches 5; Indels 0; Gaps 0;  
 QY 2 RRVYDALNVLMMNI 16  
 DB 22 REYQDLLNVKWLDI 36  
 RESULT 45  
 ABB24904  
 ID ABB24904 standard; Protein; 41 AA.  
 XX  
 AC ABB24904;  
 XX  
 XX  
 DT 23-JAN-2002 (first entry)  
 DE Protein #6903 encoded by probe for measuring heart cell gene expression.  
 XX Human; gene expression; heart; microarray; vascular system;  
 KW cardiovascular disease; hypertension; cardiac arrhythmia;  
 KW congenital heart disease.  
 XX  
 XX Homo sapiens.  
 OS  
 XX WO200157274-A2.  
 PN  
 PD 09-AUG-2001.  
 XX  
 XX 30-JAN-2001; 2001WO-US00666.  
 PF  
 XX 04-FEB-2000; 2000US-0180312.  
 PR 26-MAY-2000; 2000US-0207456.  
 PR 30-JUN-2000; 2000US-0608408.  
 PR 03-AUG-2000; 2000US-0632366.  
 PR 21-SEP-2000; 2000US-0234687.  
 PR 27-SEP-2000; 2000US-0236359.  
 PR 04-OCT-2000; 2000GB-0024263.  
 XX  
 XX (MOLE-) MOLECULAR DYNAMICS INC.  
 PA  
 XX Penn SG, Hanzel DK, Chen W, Rank DR;  
 PI  
 XX WPI; 2001-488899/53.  
 XX  
 DR Single exon nucleic acid probes for analyzing gene expression in human  
 XX hearts -  
 PT  
 XX  
 PS Claim 15; SEQ ID No 26674; 530pp; English.  
 XX  
 CC The present invention relates to single exon nucleic acid probes for  
 CC measuring human gene expression in a sample derived from human heart (see  
 CC ABA21535-ABA41305). The present sequence is a protein encoded by one such  
 CC probe. The probes may be used for predicting, measuring and displaying  
 CC gene expression in samples derived from the human heart via microarrays.  
 CC By measuring gene expression, the probes are useful for predicting,  
 CC diagnosing, grading, staging, monitoring and prognosing diseases of the  
 CC human heart and vascular system e.g. cardiovascular disease,  
 CC hypertension, cardiac arrhythmias and congenital heart disease.  
 CC Note: The sequence data for this patent did not form part of the printed  
 CC specification, but was obtained in electronic format directly from WIPO  
 CC at ftp.wipo.int/pub/published\_pct\_sequences.

XX  
SQ Sequence 41 AA;  
Query Match 39.6%; Score 36; DB 22; Length 41;  
Best Local Similarity 53.3%; Pred. No. 19;  
Matches 8; Conservative 2; Mismatches 5; Indels 0; Gaps 0;  
QY 2 RRVYDALNVLMMNI 16  
Db 22 REYQDLLNVKMALDI 36

Search completed: February 17, 2004, 10:53:41  
Job time : 37.1188 secs

**THIS PAGE BLANK (USPTO)**



GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: February 17, 2004, 10:53:48 ; Search time 27.8416 Seconds  
(without alignments)  
142.889 Million cell updates/sec

Title: US-09-900-147-3

Perfect score: 91

Sequence: 1 RRRVDALNVLMMNIISK 19

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 801455 seqs, 209382283 residues

Total number of hits satisfying chosen parameters: 318354

Minimum DB seq length: 0

Maximum DB seq length: 100

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications AA:

- 1: /cgn2\_6/ptodata/1/pubpaa/US07\_PUBCOMB.pep.\*
- 2: /cgn2\_6/ptodata/1/pubpaa/PCT\_NEW\_PUB.pep.\*
- 3: /cgn2\_6/ptodata/1/pubpaa/US06\_NEW\_PUB.pep.\*
- 4: /cgn2\_6/ptodata/1/pubpaa/US06\_PUBCOMB.pep.\*
- 5: /cgn2\_6/ptodata/1/pubpaa/US07\_NEW\_PUB.pep.\*
- 6: /cgn2\_6/ptodata/1/pubpaa/PCTUS\_PUBCOMB.pep.\*
- 7: /cgn2\_6/ptodata/1/pubpaa/US08\_NEW\_PUB.pep.\*
- 8: /cgn2\_6/ptodata/1/pubpaa/US08\_PUBCOMB.pep.\*
- 9: /cgn2\_6/ptodata/1/pubpaa/US09A\_PUBCOMB.pep.\*
- 10: /cgn2\_6/ptodata/1/pubpaa/US09B\_PUBCOMB.pep.\*
- 11: /cgn2\_6/ptodata/1/pubpaa/US09C\_PUBCOMB.pep.\*
- 12: /cgn2\_6/ptodata/1/pubpaa/US09\_NEW\_PUB.pep.\*
- 13: /cgn2\_6/ptodata/1/pubpaa/US10A\_PUBCOMB.pep.\*
- 14: /cgn2\_6/ptodata/1/pubpaa/US10B\_PUBCOMB.pep.\*
- 15: /cgn2\_6/ptodata/1/pubpaa/US10C\_PUBCOMB.pep.\*
- 16: /cgn2\_6/ptodata/1/pubpaa/US10\_NEW\_PUB.pep.\*
- 17: /cgn2\_6/ptodata/1/pubpaa/US60\_NEW\_PUB.pep.\*
- 18: /cgn2\_6/ptodata/1/pubpaa/US60\_PUBCOMB.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	91	100.0	19	10	US-09-900-147-3
2	91	100.0	37	10	US-09-900-147-1
3	91	100.0	74	15	US-10-214-188-10
4	83	91.2	19	10	US-09-900-147-15
5	79	86.8	19	10	US-09-900-147-17
6	77	84.6	19	10	US-09-900-147-16
7	76	83.5	16	10	US-09-900-147-5
8	72	79.1	30	10	US-09-900-147-6
9	69	75.8	14	10	US-09-900-147-11
10	58	63.7	96	9	US-09-864-761-45697
11	54	59.3	11	10	US-09-900-147-9
12	51	56.0	20	10	US-09-900-147-4
13	51	56.0	74	15	US-10-214-188-5
14	51	56.0	74	15	US-10-214-188-7
15	51	56.0	75	15	US-10-214-188-9

16	51	56.0	76	12	US-10-029-386-29071	Sequence 29071, A
17	51	56.0	85	11	US-09-932-581-5	Sequence 5, Appli
18	51	56.0	85	12	US-10-338-294-5	Sequence 5, Appli
19	51	56.0	85	14	US-10-165-614-2	Sequence 2, Appli
20	50	54.9	69	15	US-10-214-188-8	Sequence 8, Appli
21	50	54.9	76	11	US-09-932-581-6	Sequence 6, Appli
22	50	54.9	76	12	US-10-338-294-6	Sequence 6, Appli
23	50	54.9	76	14	US-10-165-614-3	Sequence 3, Appli
24	49	53.8	74	15	US-10-214-188-6	Sequence 2, Appli
25	42	46.2	9	10	US-09-900-147-2	Sequence 2, Appli
26	39	42.9	80	9	US-09-801-574-48	Sequence 48, Appli
27	38	41.8	48	14	US-10-001-870-180	Sequence 180, App
28	38	41.8	63	11	US-09-764-872-344	Sequence 344, App
29	36	39.6	41	9	US-09-864-761-40202	Sequence 40202, A
30	36	39.6	54	12	US-09-933-767-455	Sequence 455, App
31	36	39.6	54	15	US-10-023-282-455	Sequence 455, App
32	36	39.6	73	9	US-09-864-761-41902	Sequence 41902, A
33	36	39.6	91	9	US-09-925-302-886	Sequence 886, App
34	35	38.5	59	12	US-10-029-386-29941	Sequence 29941, A
35	35	38.5	90	9	US-09-925-299-824	Sequence 824, App
36	35	38.5	90	11	US-09-925-299-824	Sequence 824, App
37	34	37.4	59	15	US-10-036-542-148	Sequence 148, App
38	34	37.4	67	9	US-09-864-761-43263	Sequence 43263, A
39	34	37.4	88	9	US-09-939-980-513	Sequence 513, App
40	33	36.3	59	14	US-10-044-359-10	Sequence 10, App
41	33	36.3	92	10	US-09-764-846-197	Sequence 197, App
42	33	36.3	92	15	US-10-091-483-197	Sequence 197, App
43	32	35.2	50	12	US-09-864-408A-8870	Sequence 8870, Ap
44	32	35.2	58	14	US-10-044-359-2	Sequence 2, Appli
45	32	35.2	60	15	US-10-036-542-101	Sequence 101, App

ALIGNMENTS

RESULT 1  
US-09-900-147-3  
; Sequence 3, Application US/09900147  
; Patent No. US20020103121A1  
; GENERAL INFORMATION:  
; APPLICANT: La Thangue, Nicholas B  
; APPLICANT: Bandara, Lashantha R  
; TITLE OF INVENTION: Peptide antagonists of DP transcription factors  
; FILE REFERENCE: 620-67  
; CURRENT FILING DATE: 2001-07-09  
; PRIOR APPLICATION NUMBER: US/09/900,147  
; PRIOR FILING DATE: EARLIER APPLICATION NUMBER: 09/308,935  
; PRIOR APPLICATION NUMBER: EARLIER FILING DATE: 1999-05-27  
; PRIOR FILING DATE: EARLIER APPLICATION NUMBER: GB 9626589.7  
; NUMBER OF SEQ ID NOS: 18  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 3  
; LENGTH: 19  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Description of Artificial Sequence: Synthetic  
US-09-900-147-3

Query Match 100.0%; Score 91; DB 10; Length 19;  
Best Local Similarity 100.0%; Pred. No. 6.2e-09;  
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 RRRVDALNVLMMNIISK 19  
Db 1 RRRVDALNVLMMNIISK 19

RESULT 2  
US-09-900-147-1  
; Sequence 1, Application US/09900147  
; Patent No. US20020103121A1

GENERAL INFORMATION:  
; APPLICANT: La Thanque, Nicholas B  
; APPLICANT: Bandara, Lasantha R  
; TITLE OF INVENTION: Peptide antagonists of DP transcription factors  
; FILE REFERENCE: 620-67  
; CURRENT APPLICATION NUMBER: US/09/900,147  
; PRIOR FILING DATE: 2001-07-09  
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 09/308,935  
; PRIOR FILING DATE: EARLIER FILING DATE: 1999-05-27  
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: GB 9626589.7  
; PRIOR FILING DATE: EARLIER FILING DATE: 1996-12-20  
; NUMBER OF SEQ ID NOS: 18  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 1  
; LENGTH: 37  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Description of Artificial Sequence: Synthetic  
US-09-900-147-1  
  
Query Match 100.0%; Score 91; DB 10; Length 37;  
Best Local Similarity 100.0%; Pred. No. 1.3e-08;  
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
QY 1 RRRVYDALNVLMMNIISK 19  
Db 4 RRRVYDALNVLMMNIISK 22  
  
RESULT 3  
US-10-214-188-10  
; Sequence 10, Application US/10214188  
; Publication No. US20030022260A1  
; GENERAL INFORMATION:  
; APPLICANT: LA THANQUE, NICHOLAS B.  
; APPLICANT: BERNARDS, RENE  
; APPLICANT: HIJWANS, ELEANORE M.  
; TITLE OF INVENTION: TRANSCRIPTION FACTOR E2F-5  
; NUMBER OF SEQUENCES: 25  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: NIXON & VANDERHUYE P.C.  
; STREET: 1100 NORTH GLEBE ROAD  
; CITY: ARLINGTON  
; STATE: VIRGINIA  
; COUNTRY: U.S.A.  
; ZIP: 22201-4714  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/10/214,188  
; FILING DATE: 08-AUG-2002  
; CLASSIFICATION: <Unknown>  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US/08/894,139  
; FILING DATE: 13-AUG-1997  
; ATTORNEY/AGENT INFORMATION:  
; NAME: WILSON, MARY J.  
; REGISTRATION NUMBER: 32,955  
; REFERENCE/DOCKET NUMBER: 620-22  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (703) 816-4000  
; TELEFAX: (703) 816-4100  
; INFORMATION FOR SEQ ID NO: 10:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 74 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: <Unknown>  
; TOPOLOGY: linear  
; MOLECULE TYPE: peptide

SEQUENCE DESCRIPTION: SEQ ID NO: 10:  
US-10-214-188-10  
  
Query Match 100.0%; Score 91; DB 15; Length 74;  
Best Local Similarity 100.0%; Pred. No. 2.8e-08;  
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
QY 1 RRRVYDALNVLMMNIISK 19  
Db 46 RRRVYDALNVLMMNIISK 64  
  
RESULT 4  
US-09-900-147-15  
; Sequence 15, Application US/09900147  
; Patent No. US20020103121A1  
; GENERAL INFORMATION:  
; APPLICANT: La Thanque, Nicholas B  
; APPLICANT: Bandara, Lasantha R  
; TITLE OF INVENTION: Peptide antagonists of DP transcription factors  
; FILE REFERENCE: 620-67  
; CURRENT APPLICATION NUMBER: US/09/900,147  
; PRIOR FILING DATE: 2001-07-09  
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 09/308,935  
; PRIOR FILING DATE: EARLIER FILING DATE: 1999-05-27  
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: GB 9626589.7  
; PRIOR FILING DATE: EARLIER FILING DATE: 1996-12-20  
; NUMBER OF SEQ ID NOS: 18  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 15  
; LENGTH: 19  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Description of Artificial Sequence: Mutant peptide  
US-09-900-147-15  
  
Query Match 91.2%; Score 83; DB 10; Length 19;  
Best Local Similarity 89.5%; Pred. No. 1.5e-07;  
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;  
  
QY 1 RRRVYDALNVLMMNIISK 19  
Db 1 RRRVYDALNVLMMNIISK 19  
  
RESULT 5  
US-09-900-147-17  
; Sequence 17, Application US/09900147  
; Patent No. US20020103121A1  
; GENERAL INFORMATION:  
; APPLICANT: La Thanque, Nicholas B  
; APPLICANT: Bandara, Lasantha R  
; TITLE OF INVENTION: Peptide antagonists of DP transcription factors  
; FILE REFERENCE: 620-67  
; CURRENT APPLICATION NUMBER: US/09/900,147  
; PRIOR FILING DATE: 2001-07-09  
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 09/308,935  
; PRIOR FILING DATE: EARLIER FILING DATE: 1999-05-27  
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: GB 9626589.7  
; PRIOR FILING DATE: EARLIER FILING DATE: 1996-12-20  
; NUMBER OF SEQ ID NOS: 18  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 17  
; LENGTH: 19  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Description of Artificial Sequence: Mutant peptide  
US-09-900-147-17  
  
Query Match 86.8%; Score 79; DB 10; Length 19;  
Best Local Similarity 89.5%; Pred. No. 7.1e-07;

Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 RRVYDALNVLMMNIISK 19  
| | | | | | | | | | | | | | | | |  
Db 1 RRVYDARNVMMNIISK 19

## RESULT 6

US-09-900-147-16  
; Sequence 16, Application US/09900147  
; Patent No. US20020103121A1  
; GENERAL INFORMATION:  
; APPLICANT: La Thangue, Nicholas B  
; APPLICANT: Bandara, Lasantha R  
; TITLE OF INVENTION: Peptide antagonists of DP transcription factors  
; FILE REFERENCE: 620-67  
; CURRENT APPLICATION NUMBER: US/09/900,147  
; CURRENT FILING DATE: 2001-07-09  
; PRIOR APPLICATION NUMBER: 09/308,935  
; PRIOR FILING DATE: EARLIER FILING DATE: 1999-05-27  
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: GB 9626589.7  
; PRIOR FILING DATE: EARLIER FILING DATE: 1996-12-20  
; NUMBER OF SEQ ID NOS: 18  
; SOFTWARE: Patentin Ver. 2.1  
; SEQ ID NO 16  
; LENGTH: 19  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Description of Artificial Sequence: Mutant peptide  
US-09-900-147-16

Query Match 84.6%; Score 77; DB 10; Length 19;  
Best Local Similarity 89.5%; Pred. No. 1.6e-06;  
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 RRVYDALNVLMMNIISK 19  
| | | | | | | | | | | | | | | | |  
Db 1 RRVYDALNVLMMNIISK 19

## RESULT 7

US-09-900-147-5  
; Sequence 5, Application US/09900147  
; Patent No. US20020103121A1  
; GENERAL INFORMATION:  
; APPLICANT: La Thangue, Nicholas B  
; APPLICANT: Bandara, Lasantha R  
; TITLE OF INVENTION: Peptide antagonists of DP transcription factors  
; FILE REFERENCE: 620-67  
; CURRENT APPLICATION NUMBER: US/09/900,147  
; CURRENT FILING DATE: 2001-07-09  
; PRIOR APPLICATION NUMBER: 09/308,935  
; PRIOR FILING DATE: EARLIER FILING DATE: 1999-05-27  
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: GB 9626589.7  
; PRIOR FILING DATE: EARLIER FILING DATE: 1996-12-20  
; NUMBER OF SEQ ID NOS: 18  
; SOFTWARE: Patentin Ver. 2.1  
; SEQ ID NO 5  
; LENGTH: 16  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Description of Artificial Sequence: Synthetic  
US-09-900-147-5

Query Match 83.5%; Score 76; DB 10; Length 16;  
Best Local Similarity 100.0%; Pred. No. 1.9e-06;  
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 3 RRVYDALNVLMMNIIS 18  
| | | | | | | | | | | | | | | | |  
Db 1 RRVYDALNVLMMNIIS 16

## RESULT 8

US-09-900-147-6  
; Sequence 6, Application US/09900147  
; Patent No. US20020103121A1  
; GENERAL INFORMATION:  
; APPLICANT: La Thangue, Nicholas B  
; APPLICANT: Bandara, Lasantha R  
; TITLE OF INVENTION: Peptide antagonists of DP transcription factors  
; FILE REFERENCE: 620-67  
; CURRENT APPLICATION NUMBER: US/09/900,147  
; CURRENT FILING DATE: 2001-07-09  
; PRIOR APPLICATION NUMBER: 09/308,935  
; PRIOR FILING DATE: EARLIER FILING DATE: 1999-05-27  
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: GB 9626589.7  
; PRIOR FILING DATE: EARLIER FILING DATE: 1996-12-20  
; NUMBER OF SEQ ID NOS: 18  
; SOFTWARE: Patentin Ver. 2.1  
; SEQ ID NO 6  
; LENGTH: 30  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Description of Artificial Sequence: Synthetic  
US-09-900-147-6

Query Match 79.1%; Score 72; DB 10; Length 30;  
Best Local Similarity 100.0%; Pred. No. 1.9e-05;  
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 5 YDALNVLMMNIISK 19  
| | | | | | | | | | | | | | | | |  
Db 1 YDALNVLMMNIISK 15

## RESULT 9

US-09-900-147-11  
; Sequence 11, Application US/09900147  
; Patent No. US20020103121A1  
; GENERAL INFORMATION:  
; APPLICANT: La Thangue, Nicholas B  
; APPLICANT: Bandara, Lasantha R  
; TITLE OF INVENTION: Peptide antagonists of DP transcription factors  
; FILE REFERENCE: 620-67  
; CURRENT APPLICATION NUMBER: US/09/900,147  
; CURRENT FILING DATE: 2001-07-09  
; PRIOR APPLICATION NUMBER: 09/308,935  
; PRIOR FILING DATE: EARLIER FILING DATE: 1999-05-27  
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: GB 9626589.7  
; PRIOR FILING DATE: EARLIER FILING DATE: 1996-12-20  
; NUMBER OF SEQ ID NOS: 18  
; SOFTWARE: Patentin Ver. 2.1  
; SEQ ID NO 11  
; LENGTH: 14  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Description of Artificial Sequence: Synthetic  
US-09-900-147-11

Query Match 75.8%; Score 69; DB 10; Length 14;  
Best Local Similarity 100.0%; Pred. No. 2.6e-05;  
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2 RRVYDALNVLMMN 15  
| | | | | | | | | | | | | | | | |  
Db 1 RRVYDALNVLMMN 14

## RESULT 10

US-09-864-761-45697  
; Sequence 45697, Application US/09864761

Patent No. US20020048763A1  
GENERAL INFORMATION:  
APPLICANT: Penn, Sharron G.  
APPLICANT: Rank, David R.  
APPLICANT: Hanzel, David K.  
APPLICANT: Chen, Wensheng  
TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR  
FILE REFERENCE: Aomica-X-1  
CURRENT APPLICATION NUMBER: US/09/864,761  
CURRENT FILING DATE: 2001-05-23  
PRIOR APPLICATION NUMBER: US 60/180,312  
PRIOR FILING DATE: 2000-02-04  
PRIOR APPLICATION NUMBER: US 60/207,456  
PRIOR FILING DATE: 2000-05-26  
PRIOR APPLICATION NUMBER: US 09/632,366  
PRIOR FILING DATE: 2000-08-03  
PRIOR APPLICATION NUMBER: GB 24263.6  
PRIOR FILING DATE: 2000-10-04  
PRIOR APPLICATION NUMBER: US 60/236,359  
PRIOR FILING DATE: 2000-09-27  
PRIOR APPLICATION NUMBER: PCT/US01/00666  
PRIOR FILING DATE: 2001-01-30  
PRIOR APPLICATION NUMBER: PCT/US01/00667  
PRIOR FILING DATE: 2001-01-30  
PRIOR APPLICATION NUMBER: PCT/US01/00664  
PRIOR FILING DATE: 2001-01-30  
PRIOR APPLICATION NUMBER: PCT/US01/00669  
PRIOR FILING DATE: 2001-01-30  
PRIOR APPLICATION NUMBER: PCT/US01/00665  
PRIOR FILING DATE: 2001-01-30  
PRIOR APPLICATION NUMBER: PCT/US01/00668  
PRIOR FILING DATE: 2001-01-30  
PRIOR APPLICATION NUMBER: PCT/US01/00663  
PRIOR FILING DATE: 2001-01-30  
PRIOR APPLICATION NUMBER: PCT/US01/00662  
PRIOR FILING DATE: 2001-01-30  
PRIOR APPLICATION NUMBER: PCT/US01/00661  
PRIOR FILING DATE: 2001-01-30  
PRIOR APPLICATION NUMBER: PCT/US01/00670  
PRIOR FILING DATE: 2001-01-30  
PRIOR APPLICATION NUMBER: US 60/234,687  
PRIOR FILING DATE: 2000-09-21  
PRIOR APPLICATION NUMBER: US 09/608,408  
PRIOR FILING DATE: 2000-06-30  
PRIOR APPLICATION NUMBER: US 09/774,203  
PRIOR FILING DATE: 2001-01-29  
NUMBER OF SEQ ID NOS: 49117  
SOFTWARE: Annomax Sequence Listing Engine vers. 1.1  
SEQ ID NO 45697  
LENGTH: 96  
TYPE: PRT  
ORGANISM: Homo sapiens  
FEATURE:  
OTHER INFORMATION: MAP TO AC021804.3  
OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 0.89  
OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 0.99  
OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 0.9  
OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 0.92  
OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 0.88  
OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 0.75  
OTHER INFORMATION: EST HUMAN HIT: BE880658.1, EVALUATE 2.00e-51  
OTHER INFORMATION: SWISSPROT HIT: Q61501, EVALUATE 5.00e-06  
US-09-864-761-45697

Query Match 63.7%; Score 58; DB 9; Length 96;  
Best Local Similarity 47.4%; Pred. No. 0.018; Mismatches 2; Indels 0; Gaps 0;  
Matches 9; Conservative 8;

Qy 1 RRRVVDALNVLMMNIISK 19  
Db 3 RRRYDIWNLESUHLVSR 21

RESULT 11  
US-09-900-147-9  
Sequence 9, Application US/09900147  
Patent No. US20020103121A1  
GENERAL INFORMATION:  
APPLICANT: La Thangue, Nicholas B  
APPLICANT: Bandara, Lasantha R  
TITLE OF INVENTION: Peptide antagonists of DP transcription factors  
FILE REFERENCE: 620-67  
CURRENT APPLICATION NUMBER: US/09/900,147  
CURRENT FILING DATE: 2001-07-09  
PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 09/308,935  
PRIOR FILING DATE: EARLIER FILING DATE: 1999-05-27  
PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: GB 9626589.7  
PRIOR FILING DATE: EARLIER FILING DATE: 1996-12-20  
NUMBER OF SEQ ID NOS: 18  
SOFTWARE: PatentIn Ver. 2.1  
SEQ ID NO 9  
LENGTH: 11  
TYPE: PRT  
ORGANISM: Artificial Sequence  
FEATURE:  
OTHER INFORMATION: Description of Artificial Sequence: Synthetic  
US-09-900-147-9

Query Match 59.3%; Score 54; DB 10; Length 11;  
Best Local Similarity 100.0%; Pred. No. 0.0075; Mismatches 11; Conservative 0; Indels 0; Gaps 0;

Qy 1 RRRVVDALNVL 11  
Db 1 RRRVVDALNVL 11

RESULT 12  
US-09-900-147-4  
Sequence 4, Application US/09900147  
Patent No. US20020103121A1  
GENERAL INFORMATION:  
APPLICANT: La Thangue, Nicholas B  
APPLICANT: Bandara, Lasantha R  
TITLE OF INVENTION: Peptide antagonists of DP transcription factors  
FILE REFERENCE: 620-67  
CURRENT APPLICATION NUMBER: US/09/900,147  
CURRENT FILING DATE: 2001-07-09  
PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 09/308,935  
PRIOR FILING DATE: EARLIER FILING DATE: 1999-05-27  
PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: GB 9626589.7  
PRIOR FILING DATE: EARLIER FILING DATE: 1996-12-20  
NUMBER OF SEQ ID NOS: 18  
SOFTWARE: PatentIn Ver. 2.1  
SEQ ID NO 4  
LENGTH: 20  
TYPE: PRT  
ORGANISM: Artificial Sequence  
FEATURE:  
OTHER INFORMATION: Description of Artificial Sequence: Synthetic  
US-09-900-147-4

Query Match 56.0%; Score 51; DB 10; Length 20;  
Best Local Similarity 100.0%; Pred. No. 0.048; Mismatches 11; Conservative 0; Indels 0; Gaps 0;

Qy 9 NVLMMNIISK 19  
Db 1 NVLMMNIISK 11

RESULT 13  
US-10-214-188-5  
Sequence 5, Application US/10214188  
Publication No. US2003002260A1

GENERAL INFORMATION:  
APPLICANT: LA THANGUE, NICHOLAS B.  
BERNARDS, RENE  
HIJMANS, ELEANORE M.  
TITLE OF INVENTION: TRANSCRIPTION FACTOR E2F-5  
NUMBER OF SEQUENCES: 25  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: NIXON & VANDERHVE P.C.  
STREET: 1100 NORTH GLEBE ROAD  
CITY: ARLINGTON  
STATE: VIRGINIA  
COUNTRY: U.S.A.  
ZIP: 22201-4714  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/10/214,188  
FILING DATE: 08-AUG-2002  
CLASSIFICATION: <Unknown>  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US/08/894,139  
FILING DATE: 13-AUG-1997  
ATTORNEY/AGENT INFORMATION:  
NAME: WILSON, MARY J.  
REGISTRATION/DOCKET NUMBER: 620-22  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (703) 816-4000  
TELEFAX: (703) 816-4100  
INFORMATION FOR SEQ ID NO: 5:  
SEQUENCE DESCRIPTION: SEQ ID NO: 5:  
US-10-214-188-5  
Query Match 56.0%; Score 51; DB 15; Length 74;  
Best Local Similarity 47.4%; Pred. No. 0.21;  
Matches 9; Conservative 5; Mismatches 5; Indels 0; Gaps 0;  
Qy 1 RRVYDALNVLMAMNIISK 19  
Db :||:||||| :||:|  
45 KRRIYDITNVLEGIQLIAK 63  
RESULT 14  
US-10-214-188-7  
Sequence 7, Application US/10214188  
Publication No. US2003002260A1  
GENERAL INFORMATION:  
APPLICANT: LA THANGUE, NICHOLAS B.  
BERNARDS, RENE  
HIJMANS, ELEANORE M.  
TITLE OF INVENTION: TRANSCRIPTION FACTOR E2F-5  
NUMBER OF SEQUENCES: 25  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: NIXON & VANDERHVE P.C.  
STREET: 1100 NORTH GLEBE ROAD  
CITY: ARLINGTON  
STATE: VIRGINIA  
COUNTRY: U.S.A.  
ZIP: 22201-4714  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30

CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/10/214,188  
FILING DATE: 08-AUG-2002  
CLASSIFICATION: <Unknown>  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US/08/894,139  
FILING DATE: 13-AUG-1997  
ATTORNEY/AGENT INFORMATION:  
NAME: WILSON, MARY J.  
REGISTRATION/DOCKET NUMBER: 620-22  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (703) 816-4000  
TELEFAX: (703) 816-4100  
INFORMATION FOR SEQ ID NO: 7:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 74 amino acids  
TYPE: amino acid  
STRANDEDNESS: <Unknown>  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
SEQUENCE DESCRIPTION: SEQ ID NO: 7:  
US-10-214-188-7  
Query Match 56.0%; Score 51; DB 15; Length 74;  
Best Local Similarity 47.4%; Pred. No. 0.21;  
Matches 9; Conservative 5; Mismatches 5; Indels 0; Gaps 0;  
Qy 1 RRVYDALNVLMAMNIISK 19  
Db :||:||||| :||:|  
45 KRRIYDITNVLEGIHLIKK 63  
RESULT 15  
US-10-214-188-9  
Sequence 9, Application US/10214188  
Publication No. US2003002260A1  
GENERAL INFORMATION:  
APPLICANT: LA THANGUE, NICHOLAS B.  
BERNARDS, RENE  
HIJMANS, ELEANORE M.  
TITLE OF INVENTION: TRANSCRIPTION FACTOR E2F-5  
NUMBER OF SEQUENCES: 25  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: NIXON & VANDERHVE P.C.  
STREET: 1100 NORTH GLEBE ROAD  
CITY: ARLINGTON  
STATE: VIRGINIA  
COUNTRY: U.S.A.  
ZIP: 22201-4714  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/10/214,188  
FILING DATE: 08-AUG-2002  
CLASSIFICATION: <Unknown>  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US/08/894,139  
FILING DATE: 13-AUG-1997  
ATTORNEY/AGENT INFORMATION:  
NAME: WILSON, MARY J.  
REGISTRATION/DOCKET NUMBER: 620-22  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (703) 816-4000  
TELEFAX: (703) 816-4100  
INFORMATION FOR SEQ ID NO: 9:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 75 amino acids  
TYPE: amino acid

17

; SOFTWARE: FastSEQ for Windows Version 4.0

; SEQ ID NO 2

; LENGTH: 85

; TYPE: PRT

; ORGANISM: homo sapiens

US-10-165-614-2

Query Match 56.0%; Score 51; DB 14; Length 85;  
Best Local Similarity 47.4%; Pred. No. 0.24;  
Matches 9; Conservative 5; Mismatches 5; Indels 5; Gaps 0;

Qy 1 RRRYDNLVLMAMNIIISK 19

Db 55 KRRYDITNVLEGIGLIAK 73

RESULT 20

US-10-214-188-8

; Sequence 8, Application US/10214188

; Publication No. US2003002260A1

; GENERAL INFORMATION:

; APPLICANT: LA THANGUE, NICHOLAS B.

; BERNARDS, RENE

; HIMANS, ELEANORE M.

; TITLE OF INVENTION: TRANSCRIPTION FACTOR E2F-5

; NUMBER OF SEQUENCES: 25

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: NIXON & VANDERHVE P.C.

; STREET: 1100 NORTH GLEBE ROAD

; CITY: ARLINGTON

; STATE: VIRGINIA

; COUNTRY: U.S.A.

; ZIP: 22201-4714

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: PatentIn Release #1.0, Version #1.30

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/10/214,188

; FILING DATE: 08-Aug-2002

; CLASSIFICATION: <Unknown>

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: US/08/894,139

; FILING DATE: 13-AUG-1997

; ATTORNEY/AGENT INFORMATION:

; NAME: WILSON, MARY J.

; REGISTRATION NUMBER: 32,955

; REFERENCE/DOCKET NUMBER: 620-22

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: (703) 816-4000

; TELEFAX: (703) 816-4100

; INFORMATION FOR SEQ ID NO: 8:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 69 amino acids

; TYPE: amino acid

; STRANDEDNESS: <Unknown>

; TOPOLOGY: linear

; MOLECULE TYPE: peptide

; SEQUENCE DESCRIPTION: SEQ ID NO: 8:

Query Match 54.9%; Score 50; DB 15; Length 69;  
Best Local Similarity 47.4%; Pred. No. 0.29;  
Matches 9; Conservative 4; Mismatches 6; Indels 6; Gaps 0;

Qy 1 RRRYDNLVLMAMNIIISK 19

Db 40 KRRYDITNVLEGIGLIAK 58

RESULT 21

US-09-932-581-6

; Sequence 6, Application US/09932581

; Publication No. US2003005026A1

; GENERAL INFORMATION:

; APPLICANT: Andrews, William H.

; APPLICANT: Foster, Christopher A.

; APPLICANT: Fraser, Stephanie

; APPLICANT: Mohammadpour, Hamid

; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR MODULATING

; FILE OF INVENTION: TELOMERASE REVERSE TRANSCRIPTASE (TERT) EXPRESSION

; FILE REFERENCE: SIER-005

; CURRENT APPLICATION NUMBER: US/09/932,581

; CURRENT FILING DATE: 2001-08-17

; PRIOR APPLICATION NUMBER: 60/227,865

; PRIOR FILING DATE: 2000-08-24

; PRIOR APPLICATION NUMBER: 60/230,174

; PRIOR FILING DATE: 2000-09-01

; PRIOR APPLICATION NUMBER: 60/238,345

; PRIOR FILING DATE: 2000-10-05

; NUMBER OF SEQ ID NOS: 25

; SOFTWARE: FastSEQ for Windows Version 4.0

; SEQ ID NO 6

; LENGTH: 76

; TYPE: PRT

; ORGANISM: human

US-09-932-581-6

Query Match 54.9%; Score 50; DB 11; Length 76;  
Best Local Similarity 47.4%; Pred. No. 0.32;  
Matches 9; Conservative 4; Mismatches 6; Indels 6; Gaps 0;

Qy 1 RRRYDNLVLMAMNIIISK 19

Db 45 KRRYDITNVLEGIGLIAK 63

RESULT 22

US-10-338-294-6

; Sequence 6, Application US/10338294

; Publication No. US20030171326A1

; GENERAL INFORMATION:

; APPLICANT: Andrews, William H.

; APPLICANT: Foster, Christopher A.

; APPLICANT: Fraser, Stephanie

; APPLICANT: Mohammadpour, Hamid

; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR MODULATING

; FILE OF INVENTION: TELOMERASE REVERSE TRANSCRIPTASE (TERT) EXPRESSION

; FILE REFERENCE: SIER-005

; CURRENT APPLICATION NUMBER: US/10/338,294

; CURRENT FILING DATE: 2003-01-07

; PRIOR APPLICATION NUMBER: US/09/932,581

; PRIOR FILING DATE: 2001-08-17

; PRIOR APPLICATION NUMBER: 60/227,865

; PRIOR FILING DATE: 2000-08-24

; PRIOR APPLICATION NUMBER: 60/230,174

; PRIOR FILING DATE: 2000-09-01

; PRIOR APPLICATION NUMBER: 60/238,345

; PRIOR FILING DATE: 2000-10-05

; NUMBER OF SEQ ID NOS: 25

; SOFTWARE: FastSEQ for Windows Version 4.0

; SEQ ID NO 6

; LENGTH: 76

; TYPE: PRT

; ORGANISM: human

US-10-338-294-6

Query Match 54.9%; Score 50; DB 12; Length 76;  
Best Local Similarity 47.4%; Pred. No. 0.32;  
Matches 9; Conservative 4; Mismatches 6; Indels 6; Gaps 0;

Qy 1 RRRYDNLVLMAMNIIISK 19

Db 45 KRRYDITNVLEGIGLIAK 63

```

; MOLECULE TYPE: peptide
; SEQUENCE DESCRIPTION: SEQ ID NO: 6:
US-10-214-188-6

Query Match      53.8%; Score 49; DB 15; Length 74;
Best Local Similarity 47.4%; Pred. No. 0.46;
Matches 9; Conservative 4; Mismatches 6; Indels 0; Gaps 0;

QY 1 RRRVYDALNVLMMNNIIISK 19
   :||:|||||:|:|:|
Db 45 KRRIYDITNVLEGIQLIRK 63

RESULT 25
US-09-900-147-2
; Sequence 2, Application US/09500147
; Patent No. US20020103121A1
; GENERAL INFORMATION:
; APPLICANT: Bandara, Lasantha R
; TITLE OF INVENTION: Peptide antagonists of DP transcription factors
; FILE REFERENCE: 620-67
; CURRENT APPLICATION NUMBER: US/09/900,147
; CURRENT FILING DATE: 2001-07-09
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 09/308,935
; PRIOR FILING DATE: EARLIER FILING DATE: 1999-05-27
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: GB 9626589.7
; PRIOR FILING DATE: EARLIER FILING DATE: 1996-12-20
; NUMBER OF SEQ ID NOS: 18
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2
; LENGTH: 9
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Synthetic
US-09-900-147-2

Query Match      46.2%; Score 42; DB 10; Length 9;
Best Local Similarity 100.0%; Pred. No. 7e+05;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 9 NVLMAMNII 17
   |||||
Db 1 NVLMAMNII 9

RESULT 26
US-09-801-574-48
; Sequence 48, Application US/09801574
; Patent No. US20020081592A1
; GENERAL INFORMATION:
; APPLICANT: Wang, Peijing Jeremy
; APPLICANT: Page, David C.
; TITLE OF INVENTION: Reproduction-Specific Genes
; FILE REFERENCE: 0399.2007-002
; CURRENT APPLICATION NUMBER: US/09/801,574
; CURRENT FILING DATE: 2001-03-07
; PRIOR APPLICATION NUMBER: 60/187,518
; PRIOR FILING DATE: 2000-03-07
; PRIOR APPLICATION NUMBER: 60/261,557
; PRIOR FILING DATE: 2001-01-12
; NUMBER OF SEQ ID NOS: 90
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 48
; LENGTH: 80
; TYPE: PRT
; ORGANISM: Mus musculus
US-09-801-574-48

Query Match      42.9%; Score 39; DB 9; Length 80;
Best Local Similarity 53.3%; Pred. No. 26;
Matches 8; Conservative 1; Mismatches 6; Indels 0; Gaps 0;

```



OY 2 RRVYDALNVLMMNI 16  
| : | | | | |  
Db 30 RKYYSVLNVAFCNI 44

## RESULT 27

US-10-001-870-180  
; Sequence 180, Application US/10001870  
; Publication No. US20020150924A1  
; GENERAL INFORMATION:  
; APPLICANT: Salceda, Susana  
; APPLICANT: Macina, Roberto  
; APPLICANT: Recipon, Hervé  
; APPLICANT: Sun, Yongming  
; APPLICANT: Liu, Chenghua  
; TITLE OF INVENTION: Compositions and Methods Relating to Prostate Specific Genes and  
; FILE REFERENCE: DEX-0283  
; CURRENT APPLICATION NUMBER: US/10/001,870  
; CURRENT FILING DATE: 2001-11-20  
; PRIOR APPLICATION NUMBER: 60/252,189  
; PRIOR FILING DATE: 2000-11-21  
; NUMBER OF SEQ ID NOS: 217  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 180  
; LENGTH: 48  
; TYPE: PRT  
; ORGANISM: Homo sapien  
US-10-001-870-180

Query Match 41.8%; Score 38; DB 14; Length 48;  
Best Local Similarity 38.9%; Pred. No. 22;  
Matches 7; Conservative 7; Mismatches 4; Indels 0; Gaps 0;

OY 2 RRVYDALNVLMMNIISK 19  
| : | | | | : : : : |  
Db 15 RRIYSALSLYISVWLSK 32

## RESULT 28

US-09-764-872-344  
; Sequence 344, Application US/09764872  
; Publication No. US20030050231A1  
; GENERAL INFORMATION:  
; APPLICANT: Rosen et al.  
; FILE REFERENCE: PAL25  
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies  
; CURRENT APPLICATION NUMBER: US/09/764,872  
; CURRENT FILING DATE: 2001-01-17  
; Prior application data removed - consult PALM or file wrapper  
; NUMBER OF SEQ ID NOS: 957  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 344  
; LENGTH: 63  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: SITE  
; LOCATION: (3)  
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids  
; NAME/KEY: SITE  
; LOCATION: (29)  
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids  
US-09-764-872-344

Query Match 41.8%; Score 38; DB 11; Length 63;  
Best Local Similarity 56.2%; Pred. No. 30;  
Matches 9; Conservative 1; Mismatches 6; Indels 0; Gaps 0;

OY 4 VYDALNVLMMNIISK 19  
| : | | | | : : : : |  
Db 8 VYFLNPLMLKNUIFK 23

## RESULT 29

US-09-864-761-40202  
; Sequence 40202, Application US/09864761  
; Patent No. US20020048763A1  
; GENERAL INFORMATION:  
; APPLICANT: Penn, Sharron G.  
; APPLICANT: Rank, David R.  
; APPLICANT: Hanzel, David K.  
; APPLICANT: Chen, Wensheng  
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR  
; FILE REFERENCE: Aeomica-X-1  
; CURRENT APPLICATION NUMBER: US/09/864,761  
; CURRENT FILING DATE: 2001-05-23  
; PRIOR APPLICATION NUMBER: US 60/180,312  
; PRIOR FILING DATE: 2000-02-04  
; PRIOR APPLICATION NUMBER: US 60/207,456  
; PRIOR FILING DATE: 2000-05-26  
; PRIOR APPLICATION NUMBER: US 09/632,366  
; PRIOR FILING DATE: 2000-08-03  
; PRIOR APPLICATION NUMBER: GB 24263.6  
; PRIOR FILING DATE: 2000-10-04  
; PRIOR APPLICATION NUMBER: US 60/236,359  
; PRIOR FILING DATE: 2000-09-27  
; PRIOR APPLICATION NUMBER: PCT/US01/00666  
; PRIOR FILING DATE: 2001-01-30  
; PRIOR APPLICATION NUMBER: PCT/US01/00667  
; PRIOR FILING DATE: 2001-01-30  
; PRIOR APPLICATION NUMBER: PCT/US01/00664  
; PRIOR FILING DATE: 2001-01-30  
; PRIOR APPLICATION NUMBER: PCT/US01/00669  
; PRIOR FILING DATE: 2001-01-30  
; PRIOR APPLICATION NUMBER: PCT/US01/00665  
; PRIOR FILING DATE: 2001-01-30  
; PRIOR APPLICATION NUMBER: PCT/US01/00668  
; PRIOR FILING DATE: 2001-01-30  
; PRIOR APPLICATION NUMBER: PCT/US01/00663  
; PRIOR FILING DATE: 2001-01-30  
; PRIOR APPLICATION NUMBER: PCT/US01/00662  
; PRIOR FILING DATE: 2001-01-30  
; PRIOR APPLICATION NUMBER: PCT/US01/00661  
; PRIOR FILING DATE: 2001-01-30  
; PRIOR APPLICATION NUMBER: PCT/US01/00670  
; PRIOR FILING DATE: 2001-01-30  
; PRIOR APPLICATION NUMBER: US 60/234,687  
; PRIOR FILING DATE: 2000-09-21  
; PRIOR APPLICATION NUMBER: US 09/608,408  
; PRIOR FILING DATE: 2000-06-30  
; PRIOR APPLICATION NUMBER: US 09/774,203  
; PRIOR FILING DATE: 2001-01-29  
; NUMBER OF SEQ ID NOS: 49117  
; SOFTWARE: Annonax Sequence Listing Engine vers. 1.1  
; SEQ ID NO 40202  
; LENGTH: 41  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
; FEATURE:  
; OTHER INFORMATION: MAP TO AC021000.3  
; OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 1.5  
; OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 1.4  
; OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 1.7  
; OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 5.2  
; OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 1.8  
; OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 1.8  
; OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 2.6  
; OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 1.8  
; OTHER INFORMATION: EST HUMAN HIT: AV724222.1, EVALUATE 9.00e-17  
; OTHER INFORMATION: SWISSPROT HIT: P54938, EVALUATE 7.00e-18  
US-09-864-761-40202

Query Match 39.6%; Score 36; DB 9; Length 41;  
Best Local Similarity 53.3%; Pred. No. 41;

Matches 8; Conservative 2; Mismatches 5; Indels 0; Gaps 0;

QY 2 RRVYDALNVLMMNI 16  
Db 22 REYQDLLNVKWLDI 36

RESULT 30  
US-09-933-767-455  
; Sequence 455, Application US/099333767  
; Publication No. US20030181692A1  
; GENERAL INFORMATION:  
; APPLICANT: Ni et al.  
; TITLE OF INVENTION: 207 Human Secreted Proteins  
; FILE REFERENCE: P2007P2  
; CURRENT APPLICATION NUMBER: US/09/933,767  
; CURRENT FILING DATE: 2001-08-22  
; PRIOR APPLICATION NUMBER: PCT/US01/05614  
; PRIOR FILING DATE: 2001-02-21  
; PRIOR FILING DATE: 60/184,836  
; PRIOR FILING DATE: 2000-02-24  
; PRIOR APPLICATION NUMBER: 60/193,170  
; PRIOR FILING DATE: 2000-03-29  
; PRIOR APPLICATION NUMBER: 03/205,258  
; PRIOR FILING DATE: 1998-12-04  
; PRIOR APPLICATION NUMBER: PCT/US98/11422  
; PRIOR FILING DATE: 1998-06-04  
; PRIOR APPLICATION NUMBER: 60/048,885  
; PRIOR FILING DATE: 1997-06-06  
; PRIOR APPLICATION NUMBER: 60/049,375  
; PRIOR FILING DATE: 1997-06-06  
; PRIOR APPLICATION NUMBER: 60/048,881  
; PRIOR FILING DATE: 1997-06-06  
; PRIOR APPLICATION NUMBER: 60/048,880  
; PRIOR FILING DATE: 1997-06-06  
; PRIOR APPLICATION NUMBER: 60/048,896  
; PRIOR FILING DATE: 1997-06-06  
; PRIOR APPLICATION NUMBER: 60/049,020  
; PRIOR FILING DATE: 1997-06-06  
; PRIOR APPLICATION NUMBER: 60/048,876  
; PRIOR FILING DATE: 1997-06-06  
; PRIOR APPLICATION NUMBER: 60/048,895  
; PRIOR FILING DATE: 1997-06-06  
; PRIOR APPLICATION NUMBER: 60/048,884  
; PRIOR FILING DATE: 1997-06-06  
; PRIOR APPLICATION NUMBER: 60/048,894  
; PRIOR FILING DATE: 1997-06-06  
; PRIOR APPLICATION NUMBER: 60/048,971  
; PRIOR FILING DATE: 1997-06-06  
; PRIOR APPLICATION NUMBER: 60/048,964  
; PRIOR FILING DATE: 1997-06-06  
; PRIOR APPLICATION NUMBER: 60/048,882  
; PRIOR FILING DATE: 1997-06-06  
; PRIOR APPLICATION NUMBER: 60/048,899  
; PRIOR FILING DATE: 1997-06-06  
; PRIOR APPLICATION NUMBER: 60/048,893  
; PRIOR FILING DATE: 1997-06-06  
; PRIOR APPLICATION NUMBER: 60/048,900  
; PRIOR FILING DATE: 1997-06-06  
; PRIOR APPLICATION NUMBER: 60/048,901  
; PRIOR FILING DATE: 1997-06-06  
; PRIOR APPLICATION NUMBER: 60/048,892  
; PRIOR FILING DATE: 1997-06-06  
; PRIOR APPLICATION NUMBER: 60/048,915  
; PRIOR FILING DATE: 1997-06-06  
; PRIOR APPLICATION NUMBER: 60/049,019  
; PRIOR FILING DATE: 1997-06-06  
; PRIOR APPLICATION NUMBER: 60/048,970  
; PRIOR FILING DATE: 1997-06-06  
; PRIOR APPLICATION NUMBER: 60/048,972  
; PRIOR FILING DATE: 1997-06-06  
; PRIOR APPLICATION NUMBER: 60/048,916  
; PRIOR FILING DATE: 1997-06-06

PRIOR APPLICATION NUMBER: 60/049,373  
PRIOR FILING DATE: 1997-06-06  
PRIOR APPLICATION NUMBER: 60/048,875  
PRIOR FILING DATE: 1997-06-06  
PRIOR APPLICATION NUMBER: 60/049,374  
PRIOR FILING DATE: 1997-06-06  
PRIOR APPLICATION NUMBER: 60/048,917  
PRIOR FILING DATE: 1997-06-06  
PRIOR APPLICATION NUMBER: 60/048,949  
PRIOR FILING DATE: 1997-06-06  
PRIOR APPLICATION NUMBER: 60/048,974  
PRIOR FILING DATE: 1997-06-06  
PRIOR APPLICATION NUMBER: 60/048,883  
PRIOR FILING DATE: 1997-06-06  
PRIOR APPLICATION NUMBER: 60/048,897  
PRIOR FILING DATE: 1997-06-06  
PRIOR APPLICATION NUMBER: 60/048,898  
PRIOR FILING DATE: 1997-06-06  
PRIOR APPLICATION NUMBER: 60/048,962  
PRIOR FILING DATE: 1997-06-06  
PRIOR APPLICATION NUMBER: 60/048,963  
PRIOR FILING DATE: 1997-06-06  
PRIOR APPLICATION NUMBER: 60/048,877  
PRIOR FILING DATE: 1997-06-06  
PRIOR APPLICATION NUMBER: 60/048,878  
PRIOR FILING DATE: 1997-06-06  
PRIOR APPLICATION NUMBER: 60/068,054  
PRIOR FILING DATE: 1997-12-18  
PRIOR APPLICATION NUMBER: 60/068,064  
PRIOR FILING DATE: 1997-12-18  
PRIOR APPLICATION NUMBER: 60/068,053  
PRIOR FILING DATE: 1997-12-18  
PRIOR APPLICATION NUMBER: 60/070,923  
PRIOR FILING DATE: 1997-12-18  
PRIOR APPLICATION NUMBER: 60/073,160  
PRIOR FILING DATE: 1998-01-30  
PRIOR APPLICATION NUMBER: 60/073,159  
PRIOR FILING DATE: 1998-01-30  
PRIOR APPLICATION NUMBER: 60/073,165  
PRIOR FILING DATE: 1998-01-30  
PRIOR APPLICATION NUMBER: 60/073,164  
PRIOR FILING DATE: 1998-01-30  
PRIOR APPLICATION NUMBER: 60/085,925  
PRIOR FILING DATE: 1998-05-18  
PRIOR APPLICATION NUMBER: 60/085,921  
PRIOR FILING DATE: 1998-05-18  
PRIOR APPLICATION NUMBER: 60/085,923  
PRIOR FILING DATE: 1998-05-18  
PRIOR APPLICATION NUMBER: 60/085,922  
PRIOR FILING DATE: 1998-05-18  
PRIOR APPLICATION NUMBER: 60/092,921  
PRIOR FILING DATE: 1998-07-15  
PRIOR APPLICATION NUMBER: 60/094,657  
PRIOR FILING DATE: 1998-07-30  
NUMBER OF SEQ ID NOS: 1245  
SOFTWARE: PatentIn Ver. 2.0  
SEQ ID NO 455  
LENGTH: 54  
TYPE: PRT  
ORGANISM: Homo sapiens  
US-09-933-767-455

Query Match 39.6%; Score 36; DB 12; Length 54;  
Best Local Similarity 41.2%; Pred. No. 55;  
Matches 7; Conservative 5; Mismatches 5; Indels 0; Gaps 0;

Qy 3 RRVYDALNVLMMNIISK 19  
Db 35 RRVYDALNVLMMNIISK 51

RESULT 31  
US-10-023-282-455

; Sequence 455, Application US/10023282  
; Publication No. US2003002893A1  
; GENERAL INFORMATION:  
; APPLICANT: Young et al.  
; TITLE OF INVENTION: 207 Human Secreted Proteins  
; FILE REFERENCE: P2007P1  
; CURRENT APPLICATION NUMBER: US/10/023,282  
; CURRENT FILING DATE: 2001-12-20  
; EARLIER APPLICATION NUMBER: 09/205,258  
; EARLIER FILING DATE: 1998-12-04  
; EARLIER APPLICATION NUMBER: PCT/US98/11422  
; EARLIER FILING DATE: 1998-06-04  
; EARLIER APPLICATION NUMBER: 60/048,885  
; EARLIER FILING DATE: 1997-06-06  
; EARLIER APPLICATION NUMBER: 60/049,375  
; EARLIER FILING DATE: 1997-06-06  
; EARLIER APPLICATION NUMBER: 60/048,881  
; EARLIER FILING DATE: 1997-06-06  
; EARLIER APPLICATION NUMBER: 60/048,880  
; EARLIER FILING DATE: 1997-06-06  
; EARLIER APPLICATION NUMBER: 60/048,896  
; EARLIER FILING DATE: 1997-06-06  
; EARLIER APPLICATION NUMBER: 60/049,020  
; EARLIER FILING DATE: 1997-06-06  
; EARLIER APPLICATION NUMBER: 60/048,876  
; EARLIER FILING DATE: 1997-06-06  
; EARLIER APPLICATION NUMBER: 60/048,895  
; EARLIER FILING DATE: 1997-06-06  
; EARLIER APPLICATION NUMBER: 60/048,884  
; EARLIER FILING DATE: 1997-06-06  
; EARLIER APPLICATION NUMBER: 60/048,964  
; EARLIER FILING DATE: 1997-06-06  
; EARLIER APPLICATION NUMBER: 60/048,882  
; EARLIER FILING DATE: 1997-06-06  
; EARLIER APPLICATION NUMBER: 60/048,899  
; EARLIER FILING DATE: 1997-06-06  
; EARLIER APPLICATION NUMBER: 60/048,893  
; EARLIER FILING DATE: 1997-06-06  
; EARLIER APPLICATION NUMBER: 60/048,900  
; EARLIER FILING DATE: 1997-06-06  
; EARLIER APPLICATION NUMBER: 60/048,901  
; EARLIER FILING DATE: 1997-06-06  
; EARLIER APPLICATION NUMBER: 60/048,892  
; EARLIER FILING DATE: 1997-06-06  
; EARLIER APPLICATION NUMBER: 60/048,915  
; EARLIER FILING DATE: 1997-06-06  
; EARLIER APPLICATION NUMBER: 60/049,019  
; EARLIER FILING DATE: 1997-06-06  
; EARLIER APPLICATION NUMBER: 60/048,970  
; EARLIER FILING DATE: 1997-06-06  
; EARLIER APPLICATION NUMBER: 60/048,972  
; EARLIER FILING DATE: 1997-06-06  
; EARLIER APPLICATION NUMBER: 60/048,916  
; EARLIER FILING DATE: 1997-06-06  
; EARLIER APPLICATION NUMBER: 60/049,373  
; EARLIER FILING DATE: 1997-06-06  
; EARLIER APPLICATION NUMBER: 60/048,875  
; EARLIER FILING DATE: 1997-06-06  
; EARLIER APPLICATION NUMBER: 60/049,374  
; EARLIER FILING DATE: 1997-06-06  
; EARLIER APPLICATION NUMBER: 60/048,917  
; EARLIER FILING DATE: 1997-06-06  
; EARLIER APPLICATION NUMBER: 60/048,949  
; EARLIER FILING DATE: 1997-06-06  
; EARLIER APPLICATION NUMBER: 60/048,974  
; EARLIER FILING DATE: 1997-06-06  
; EARLIER APPLICATION NUMBER: 60/048,883  
; EARLIER FILING DATE: 1997-06-06  
; EARLIER APPLICATION NUMBER: 60/048,897

; EARLIER FILING DATE: 1997-06-06  
; EARLIER APPLICATION NUMBER: 60/048,898  
; EARLIER FILING DATE: 1997-06-06  
; EARLIER APPLICATION NUMBER: 60/048,962  
; EARLIER FILING DATE: 1997-06-06  
; EARLIER APPLICATION NUMBER: 60/048,963  
; EARLIER FILING DATE: 1997-06-06  
; EARLIER APPLICATION NUMBER: 60/048,877  
; EARLIER FILING DATE: 1997-06-06  
; EARLIER APPLICATION NUMBER: 60/048,878  
; EARLIER FILING DATE: 1997-06-06  
; EARLIER APPLICATION NUMBER: 60/070,923  
; EARLIER FILING DATE: 1997-12-18  
; EARLIER APPLICATION NUMBER: 60/092,921  
; EARLIER FILING DATE: 1998-07-15  
; EARLIER APPLICATION NUMBER: 60/094,657  
; EARLIER FILING DATE: 1998-07-30  
; NUMBER OF SEQ ID NOS: 1227  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 455  
; LENGTH: 54  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-10-023-282-455  
  
Query Match 39.6%; Score 36; DB 15; Length 54;  
Best Local Similarity 41.2%; Pred. No. 55;  
Matches 7; Conservative 5; Mismatches 5; Indels 0; Gaps 0;  
  
QY 3 RYVDALNVLMMNIISK 19  
||| |::: |::| |  
Db 35 RYVPAMHFTLCVHIYSK 51  
  
RESULT 32  
US-09-864-761-41902  
; Sequence 41902, Application US/09864761  
; Patent No. US20020048763A1  
; GENERAL INFORMATION:  
; APPLICANT: Penn, Sharron G.  
; APPLICANT: Hanzel, David K.  
; APPLICANT: Chen, Wensheng  
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR  
; FILE REFERENCE: Aeomica-X-1  
; CURRENT APPLICATION NUMBER: US/09/864,761  
; CURRENT FILING DATE: 2001-05-23  
; PRIOR APPLICATION NUMBER: US 60/180,312  
; PRIOR FILING DATE: 2000-02-04  
; PRIOR APPLICATION NUMBER: US 60/207,456  
; PRIOR FILING DATE: 2000-05-26  
; PRIOR APPLICATION NUMBER: US 09/632,366  
; PRIOR FILING DATE: 2000-08-03  
; PRIOR APPLICATION NUMBER: GB 24263.6  
; PRIOR FILING DATE: 2000-10-04  
; PRIOR APPLICATION NUMBER: US 60/236,359  
; PRIOR FILING DATE: 2000-09-27  
; PRIOR APPLICATION NUMBER: PCT/US01/00666  
; PRIOR FILING DATE: 2001-01-30  
; PRIOR APPLICATION NUMBER: PCT/US01/00667  
; PRIOR FILING DATE: 2001-01-30  
; PRIOR APPLICATION NUMBER: PCT/US01/00664  
; PRIOR FILING DATE: 2001-01-30  
; PRIOR APPLICATION NUMBER: PCT/US01/00669  
; PRIOR FILING DATE: 2001-01-30  
; PRIOR APPLICATION NUMBER: PCT/US01/00665  
; PRIOR FILING DATE: 2001-01-30  
; PRIOR APPLICATION NUMBER: PCT/US01/00668  
; PRIOR FILING DATE: 2001-01-30  
; PRIOR APPLICATION NUMBER: PCT/US01/00663  
; PRIOR FILING DATE: 2001-01-30  
; PRIOR APPLICATION NUMBER: PCT/US01/00662



US-09-925-299-824  
; Sequence 824, Application US/09925299  
; Publication No. US20030040617A9  
; GENERAL INFORMATION:  
; APPLICANT: Rosen et al.  
; TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies  
; FILE REFERENCE: PA102  
; CURRENT APPLICATION NUMBER: US/09/925,299  
; CURRENT FILING DATE: 2001-08-10  
; PRIOR APPLICATION NUMBER: PCT/US00/05883  
; PRIOR FILING DATE: 2000-03-08  
; PRIOR APPLICATION NUMBER: 60/124,270  
; PRIOR FILING DATE: 1999-03-12  
; NUMBER OF SEQ ID NOS: 1556  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 824  
; LENGTH: 90  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: SITE  
; LOCATION: (36)  
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids  
US-09-925-299-824

Query Match 38.5%; Score 35; DB 11; Length 90;  
Best Local Similarity 75.0%; Pred. No. 1.5e+02;  
Matches 6; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Qy 5 YDALNVLM 12  
:|||||:  
Db 42 FDALNTLM 49

RESULT 37  
US-10-036-542-148  
; Sequence 148, Application US/10036542  
; Publication No. US20030083481A1  
; GENERAL INFORMATION:  
; APPLICANT: Birse et al.  
; TITLE OF INVENTION: 25 Human Prostate and Prostate Cancer Associated Proteins  
; FILE REFERENCE: PA002P1  
; CURRENT APPLICATION NUMBER: US/10/036,542  
; CURRENT FILING DATE: 2002-01-07  
; PRIOR APPLICATION NUMBER: PCT/US00/19666  
; PRIOR FILING DATE: 2000-07-20  
; PRIOR APPLICATION NUMBER: 60/144,972  
; PRIOR FILING DATE: 1999-07-21  
; PRIOR APPLICATION NUMBER: 60/148,681  
; PRIOR FILING DATE: 1999-08-13  
; PRIOR APPLICATION NUMBER: 60/149,173  
; PRIOR FILING DATE: 1999-08-17  
; PRIOR APPLICATION NUMBER: 60/158,004  
; PRIOR FILING DATE: 1999-10-06  
; PRIOR APPLICATION NUMBER: 60/194,689  
; PRIOR FILING DATE: 2000-04-05  
; NUMBER OF SEQ ID NOS: 157  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 148  
; LENGTH: 59  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-10-036-542-148

Query Match 37.4%; Score 34; DB 15; Length 59;  
Best Local Similarity 50.0%; Pred. No. 1.3e+02;  
Matches 6; Conservative 5; Mismatches 1; Indels 0; Gaps 0;

Qy 3 RVDALNVLMAM 14  
|||:|:|:  
Db 38 RVEYLNLLISV 49

RESULT 38  
US-09-864-761-43263  
; Sequence 43263, Application US/09864761  
; Patent No. US20020048763A1  
; GENERAL INFORMATION:  
; APPLICANT: Penn, Sharron G.  
; APPLICANT: Rank, David R.  
; APPLICANT: Hanzel, David K.  
; APPLICANT: Chen, Weisheng  
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR  
; FILE REFERENCE: Aeomica-X-1  
; CURRENT APPLICATION NUMBER: US/09/864,761  
; CURRENT FILING DATE: 2001-05-23  
; PRIOR APPLICATION NUMBER: US 60/180,312  
; PRIOR FILING DATE: 2000-02-04  
; PRIOR APPLICATION NUMBER: US 60/207,456  
; PRIOR FILING DATE: 2000-05-26  
; PRIOR APPLICATION NUMBER: US 09/632,366  
; PRIOR FILING DATE: 2000-08-03  
; PRIOR APPLICATION NUMBER: GB 24263.6  
; PRIOR FILING DATE: 2000-10-04  
; PRIOR APPLICATION NUMBER: US 60/236,359  
; PRIOR FILING DATE: 2000-09-27  
; PRIOR APPLICATION NUMBER: PCT/US01/00666  
; PRIOR FILING DATE: 2001-01-30  
; PRIOR APPLICATION NUMBER: PCT/US01/00667  
; PRIOR FILING DATE: 2001-01-30  
; PRIOR APPLICATION NUMBER: PCT/US01/00664  
; PRIOR FILING DATE: 2001-01-30  
; PRIOR APPLICATION NUMBER: PCT/US01/00669  
; PRIOR FILING DATE: 2001-01-30  
; PRIOR APPLICATION NUMBER: PCT/US01/00665  
; PRIOR FILING DATE: 2001-01-30  
; PRIOR APPLICATION NUMBER: PCT/US01/00668  
; PRIOR FILING DATE: 2001-01-30  
; PRIOR APPLICATION NUMBER: PCT/US01/00663  
; PRIOR FILING DATE: 2001-01-30  
; PRIOR APPLICATION NUMBER: PCT/US01/00662  
; PRIOR FILING DATE: 2001-01-30  
; PRIOR APPLICATION NUMBER: PCT/US01/00661  
; PRIOR FILING DATE: 2001-01-30  
; PRIOR APPLICATION NUMBER: PCT/US01/00670  
; PRIOR FILING DATE: 2001-01-30  
; PRIOR APPLICATION NUMBER: US 60/234,687  
; PRIOR FILING DATE: 2000-09-21  
; PRIOR APPLICATION NUMBER: US 09/608,408  
; PRIOR FILING DATE: 2000-06-30  
; PRIOR APPLICATION NUMBER: US 09/774,203  
; PRIOR FILING DATE: 2001-01-29  
; NUMBER OF SEQ ID NOS: 49117  
; SOFTWARE: Annonax Sequence Listing Engine vers. 1.1  
; SEQ ID NO 43263  
; LENGTH: 67  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
; FEATURE:  
; OTHER INFORMATION: MAP TO AL031301.1  
; OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 1.6  
US-09-864-761-43263

Query Match 37.4%; Score 34; DB 9; Length 67;  
Best Local Similarity 35.7%; Pred. No. 1.6e+02;  
Matches 5; Conservative 6; Mismatches 3; Indels 0; Gaps 0;

Qy 4 YVDALNVLMAMNII 17  
:|:|:|:  
Db 38 IITINIINIINI 51

RESULT 39  
US-09-939-980-513  
; Sequence 513, Application US/09939980

Patent No. US20020082234A1  
GENERAL INFORMATION:  
APPLICANT: Black, Michael  
Burnham, Martin  
Hodgson, John  
Knowles, David  
Lonetto, Michael  
Nicholas, Richard  
Pratt, Julie  
Reichard, Richard  
Rosenberg, Martin  
Ward, Judith  
TITLE OF INVENTION: No. US20020082234A1el Prokaryotic Polynucleotides,  
Polypeptides and Their Uses  
NUMBER OF SEQUENCES: 534  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: SmithKline Beecham Corporation  
STREET: 709 Swedeland Road  
CITY: King of Prussia  
STATE: PA  
COUNTRY: USA  
ZIP: 19406-0939  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette  
COMPUTER: IBM Compatible  
OPERATING SYSTEM: DOS  
SOFTWARE: FASSEQ for Windows Version 2.0  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/939,980  
FILING DATE: 27-Aug-2001  
CLASSIFICATION: <Unknown>  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/936,165  
FILING DATE: <Unknown>  
ATTORNEY/AGENT INFORMATION:  
NAME: Gimmi, Edward R  
REGISTRATION NUMBER: 38,891  
REFERENCE/DOCKET NUMBER: P50549  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 610-270-4478  
TELEFAX: 610-270-5090  
TELEX: <Unknown>  
INFORMATION FOR SEQ ID NO: 513:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 88 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: Protein  
SEQUENCE DESCRIPTION: SEQ ID NO: 513:  
US-09-939-980-513  
Query Match 37.4%; Score 34; DB 9; Length 88;  
Best Local Similarity 56.2%; Pred. No. 2.1e+02;  
Matches 9; Conservative 3; Mismatches 2; Indels 2; Gaps 1;  
QY 4 VYDAL--NVLMAMNII 17  
DB 48 MYDAKMDNVLPINII 63  
RESULT 40  
US-10-044-359-10  
Sequence 10, Application US/10044359  
Publication No. US20020160454A1  
GENERAL INFORMATION:  
APPLICANT: Herrman, Rafael  
APPLICANT: Wong, James F.  
APPLICANT: Lee, Jian-Ming  
TITLE OF INVENTION: SCORPION TOXINS  
FILE REFERENCE: BBI367 US NA  
CURRENT APPLICATION NUMBER: US/10/044,359  
CURRENT FILING DATE: 2002-01-11

PRIOR APPLICATION NUMBER: 09/599,416  
PRIOR FILING DATE: 2000-06-22  
PRIOR APPLICATION NUMBER: 60/140,227  
PRIOR FILING DATE: 1999-06-22  
NUMBER OF SEQ ID NOS: 28  
SOFTWARE: Microsoft Office 97  
SEQ ID NO 10  
LENGTH: 59  
TYPE: PRT  
ORGANISM: Hottentotta judaica  
US-10-044-359-10  
Query Match 36.3%; Score 33; DB 14; Length 59;  
Best Local Similarity 33.3%; Pred. No. 2e+02;  
Matches 5; Conservative 6; Mismatches 4; Indels 0; Gaps 0;  
QY 3 RYVDALNVLMAMNII 17  
DB 5 RLYAILIIVLMNVI 19  
RESULT 41  
US-09-764-846-197  
Sequence 197, Application US/09764846  
Patent No. US20020102638A1  
GENERAL INFORMATION:  
APPLICANT: Rosen et al.  
TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies  
FILE REFERENCE: PT212  
CURRENT APPLICATION NUMBER: US/09/764,846  
CURRENT FILING DATE: 2001-01-17  
Prior application data removed - consult PALM or file wrapper  
NUMBER OF SEQ ID NOS: 348  
SOFTWARE: PatentIn Ver. 2.0  
SEQ ID NO 197  
LENGTH: 92  
TYPE: PRT  
ORGANISM: Homo sapiens  
FEATURE:  
NAME/KEY: SITE  
LOCATION: (88)  
OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids  
US-09-764-846-197  
Query Match 36.3%; Score 33; DB 10; Length 92;  
Best Local Similarity 50.0%; Pred. No. 3.3e+02;  
Matches 6; Conservative 4; Mismatches 2; Indels 0; Gaps 0;  
QY 8 LNVLMAMNIIISK 19  
DB 72 LSVLLCKNVITK 83  
RESULT 42  
US-10-091-483-197  
Sequence 197, Application US/10091483  
Publication No. US20030049650A1  
GENERAL INFORMATION:  
APPLICANT: Rosen et al.  
TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies  
FILE REFERENCE: PT212C1  
CURRENT APPLICATION NUMBER: US/10/091,483  
CURRENT FILING DATE: 2002-03-07  
NUMBER OF SEQ ID NOS: 348  
Prior Application removed - See File Wrapper or Palm  
SOFTWARE: PatentIn Ver. 2.0  
SEQ ID NO 157  
LENGTH: 92  
TYPE: PRT  
ORGANISM: Homo sapiens  
FEATURE:  
NAME/KEY: misc\_feature  
LOCATION: (88)

```
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
US-10-091-483-197

Query Match      36.3%; Score 33; DB 15; Length 92;
Best Local Similarity 50.0%; Pred. No. 3.3e+02;
Matches 6; Conservative 4; Mismatches 2; Indels 0; Gaps 0;

QY      8 LNVLMAMNIIISK 19
       |:|:|:|:|:|
Db      72 LSVLLCKNVITK 83

RESULT 43
US-09-864-408A-8870
; Sequence 8870, Application US/09864408A
; Publication No. US20040009474A1
; GENERAL INFORMATION:
; APPLICANT: Leach, Martin D.
; APPLICANT: Shimkets, Richard A.
; TITLE OF INVENTION: No. US20040009474A1 Human Polynucleotides and Polypeptides Encod
; FILE REFERENCE: 21402-012
; CURRENT APPLICATION NUMBER: US/09/864,408A
; CURRENT FILING DATE: 2001-05-24
; PRIOR APPLICATION NUMBER: 60/206,690
; PRIOR FILING DATE: 2000-05-24
; NUMBER OF SEQ ID NOS: 9068
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 8870
; LENGTH: 50
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-864-408A-8870

Query Match      35.2%; Score 32; DB 12; Length 50;
Best Local Similarity 54.5%; Pred. No. 2.5e+02;
Matches 6; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY      1 RRVYDALNVIL 11
       |:|:|:|:|
Db      9 RQLHDARNAL 19

RESULT 44
US-10-044-359-2
; Sequence 2, Application US/10044359
; Publication No. US20020160454A1
; GENERAL INFORMATION:
; APPLICANT: Heriman, Rafael
; APPLICANT: Wong, James F.
; APPLICANT: Lee, Jian-Ming
; TITLE OF INVENTION: SCORPION TOXINS
; FILE REFERENCE: B1367 US NA
; CURRENT APPLICATION NUMBER: US/10/044,359
; CURRENT FILING DATE: 2002-01-11
; PRIOR APPLICATION NUMBER: 09/599,416
; PRIOR FILING DATE: 2000-06-22
; PRIOR APPLICATION NUMBER: 60/140,227
; PRIOR FILING DATE: 1999-06-22
; NUMBER OF SEQ ID NOS: 28
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 2
; LENGTH: 58
; TYPE: PRT
; ORGANISM: Hottentotta judaica
US-10-044-359-2

Query Match      35.2%; Score 32; DB 14; Length 58;
Best Local Similarity 33.3%; Pred. No. 2.9e+02;
Matches 5; Conservative 6; Mismatches 4; Indels 0; Gaps 0;

QY      3 RVDALNVLMAMNII 17
       |:|:|:|:|:|
Db      3 RIETIILIVFALNII 17
```

```
RESULT 45
US-10-036-542-101
; Sequence 101, Application US/10036542
; Publication No. US20030083481A1
; GENERAL INFORMATION:
; APPLICANT: Birse et al.
; TITLE OF INVENTION: 25 Human Prostate and Prostate Cancer Associated Proteins
; FILE REFERENCE: PA002P1
; CURRENT APPLICATION NUMBER: US/10/036,542
; CURRENT FILING DATE: 2002-01-07
; PRIOR APPLICATION NUMBER: PCT/US00/19666
; PRIOR FILING DATE: 2000-07-20
; PRIOR APPLICATION NUMBER: 60/144,972
; PRIOR FILING DATE: 1999-07-21
; PRIOR APPLICATION NUMBER: 60/148,681
; PRIOR FILING DATE: 1999-08-13
; PRIOR APPLICATION NUMBER: 60/149,173
; PRIOR FILING DATE: 1999-08-17
; PRIOR APPLICATION NUMBER: 60/158,004
; PRIOR FILING DATE: 1999-10-06
; PRIOR APPLICATION NUMBER: 60/194,689
; PRIOR FILING DATE: 2000-04-05
; NUMBER OF SEQ ID NOS: 157
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 101
; LENGTH: 60
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-036-542-101
```

```
Query Match      35.2%; Score 32; DB 15; Length 60;
Best Local Similarity 46.2%; Pred. No. 3e+02;
Matches 6; Conservative 5; Mismatches 2; Indels 0; Gaps 0;

QY      2 RRVYDALNVLMAM 14
       |:|:|:|:|:|
Db      38 RDVVEVNLNLIISV 50
```

Search completed: February 17, 2004, 11:02:20  
Job time : 27.8416 secs

**THIS PAGE BLANK (USPTO)**





C.; Shiba, T.; Hattori, M.; Ogasawara, N.; Hayashi, H.; Hiramatsu, K.  
Lancet 357, 1225-1240, 2001  
A>Title: Whole genome sequencing of methicillin-resistant *Staphylococcus aureus*.  
A;Reference number: AB9758; MUID:21311952; PMID:11418146  
A;Accession: D89831  
A>Status: preliminary  
A:Molecule type: DNA  
A;Residues: 1-74 <KUR>  
A;Cross-references: GB:BA000018; PID:g13700510; PIDN:BA041807.1; GSPDB:GN00149  
A;Experimental source: strain N315  
C;Genetics:  
A;Gene: SA0575

Query Match 37.4%; Score 34; DB 2; Length 74;  
Best Local Similarity 50.0%; Pred. No. 85;  
Matches 6; Conservative 4; Mismatches 2; Indels 0; Gaps 0

QY 6 DALNVLMMWII 17  
|:| | :|||  
DB 35 DSMALVDLII 46

RESULT 6  
C69425  
transcription regulator asnC - Archaeoglobus fulgidus  
C;Species: Archaeoglobus fulgidus  
C;Date: 10-Sep-1999 #sequence\_revision 10-Sep-1999 #text\_change 21-Jul-2000  
C;Accession: C69425  
R;Klenk, H.P.; Clayton, R.A.; Tomb, J.P.; White, O.; Nelson, K.E.; Ketchum, K.A.; Fleischmann, R.D.; Quackenbush, J.; Lee, N.H.; Sutton, G.G.; Gill, S.; Kirk, G.; Glodek, A.; Zhou, L.; Overbeek, R.; Gocayne, J.D.; Weidman, J.F.; McDonald, L. Nature 390, 364-370, 1997  
A;Authors: Usterback, T.; Cotton, M.D.; Spriggs, T.; Artiach, P.; Kaine, B.P.; Smith, H.O.; Woese, C.R.; Venter, J.C.  
A;Title: The complete genome sequence of the hyperthermophilic, sulfate-reducing bacterium *Archaeoglobus fulgidus*.  
A;Reference number: A69250; MUID:98049343; PMID:9389475  
A;Accession: C69425  
A>Status: preliminary; nucleic acid sequence not shown; translation not shown  
A:Molecule type: DNA  
A;Residues: 1-77 <KLE>  
A;Cross-references: GB:AE001007; GB:AE000782; NID:92689330; PIDN:AA089844.1; F000000000  
C;Superfamily: Archaeoglobus fulgidus transcription regulator asnC

Query Match 37.4%; Score 34; DB 1; Length 77;  
Best Local Similarity 70.0%; Pred. No. 89;  
Matches 7; Conservative 1; Mismatches 2; Indels 0; Gaps 0

QY 2 RRVDALNVL 11  
|:| | | | |  
DB 16 RKVDALAL 25

RESULT 7  
T42312  
hypothetical protein - phage SPPI  
C;Species: phage SPPI  
C;Date: 03-Dec-1999 #sequence\_revision 03-Dec-1999 #text\_change 11-May-2000  
C;Accession: T42312  
R;Alonso, J.C.; Luder, G.; Stiege, A.C.; Chai, S.; Weise, F.; Trautner, T.A. Gene 204, 203-212, 1997  
A;Title: The complete nucleotide sequence and functional organization of bacteriophage SPPI.  
A;Reference number: Z22137; MUID:98094274; PMID:9434185  
A;Accession: T42312  
A>Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: DNA  
A;Residues: 1-82 <ALO>  
A;Cross-references: EMBL:X97918; PIDN:CAA66519.1

Query Match 37.4%; Score 34; DB 2; Length 82;  
Best Local Similarity 46.2%; Pred. No. 95;  
Matches 6; Conservative 5; Mismatches 2; Indels 0; Gaps 0

QY 2 RRVYDALNVLAM 14



A;Status: preliminary; nucleic acid sequence not shown; translation not shown

A;Molecule type: DNA

A;Residues: 1-93 <AQF>

A;Cross-references: GB:AE000727; NID:g2983623; PIDN:AAC07208.1; PID:g2983640; GB:AE00065;

A;Experimental source: strain VP5

C;Genetics:

A;Gene: flie

C;Superfamily: flagellar hook-basal body protein flie

Query Match 35.2%; Score 32; DB 2; Length 93;

Best Local Similarity 37.5%; Pred. No. 2.4e+02; Indels 0; Gaps 0;

Matches 6; Conservative 4; Mismatches 6;

QY 1 RRRVYDALNVLMMNI 16

DB 78 RNKLEAYNELMKQV 93

RESULT 16

H84227

hypothetical protein Vng0703h [imported] - Halobacterium sp. NRC-1

C;Species: Halobacterium sp. NRC-1

C;Date: 02-Feb-2001 #sequence\_revision 02-Feb-2001 #text\_change 02-Feb-2001

C;Accession: H84227

R;Ng, W.V.; Kennedy, S.P.; Mahairas, G.G.; Berquist, B.; Pan, M.; Shukla, H.D.; Laaky, S.; Leithauser, B.; Keller, K.; Cruz, R.; Danson, M.J.; Hough, D.W.; Maddocks, D.G.; Jablon, J.; Jung, K.H.; Alam, M.; Freitas, T.

Proc. Natl. Acad. Sci. U.S.A. 97, 12176-12181, 2000

A;Authors: Hou, S.; Daniels, C.J.; Dennis, P.P.; Omer, A.D.; Ebhardt, H.; Lowe, T.M.; Li;

A;Title: Genome sequence of Halobacterium species NRC-1

A;Reference number: A84160; MUID:20504483; PMID:11016950

A;Accession: H84227

A;Status: preliminary

A;Molecule type: DNA

A;Residues: 1-94 <STO>

A;Cross-references: GB:AE004437; NID:gl0580284; PIDN:AAG19188.1; GSPDB:GN00138

C;Genetics:

A;Gene: VNG0703H

Query Match 35.2%; Score 32; DB 2; Length 94;

Best Local Similarity 36.8%; Pred. No. 2.5e+02;

Matches 7; Conservative 5; Mismatches 7; Indels 0; Gaps 0;

QY 1 RRRVYDALNVLMMNIISK 19

DB 41 QRTVRDALDRLQADVVEK 59

RESULT 17

E69550

conserved hypothetical protein AF2404 - Archaeoglobus fulgidus

C;Species: Archaeoglobus fulgidus

C;Date: 05-Dec-1997 #sequence\_revision 05-Dec-1997 #text\_change 08-Oct-1999

C;Accession: E69550

R;Klenk, H.P.; Clayton, R.A.; Tomb, J.F.; White, O.; Nelson, K.E.; Ketchum, K.A.; Dodson, J.; Fleischmann, R.D.; Quackenbush, J.; Lee, N.H.; Sutton, G.G.; Gill, S.; Kirkness, E.F.; Glodek, A.; Zhou, L.; Overbeek, R.; Gocayne, J.D.; Weidman, J.F.; McDonald, L. Nature 390, 364-370, 1997

A;Authors: Usterback, T.; Cotton, M.D.; Spriggs, T.; Artiach, P.; Kaine, B.P.; Sykes, S.N. Smith, H.O.; Woese, C.R.; Venter, J.C.

A;Title: The complete genome sequence of the hyperthermophilic, sulfate-reducing archaeon

A;Reference number: A69250; MUID:98049343; PMID:9389475

A;Accession: E69550

A;Status: preliminary; nucleic acid sequence not shown; translation not shown

A;Molecule type: DNA

A;Residues: 1-96 <KLE>

A;Cross-references: GB:AE001110; GB:AE000782; NID:g2689433; PIDN:AA91262.1; PID:g2650691

C;Superfamily: conserved hypothetical protein AF2404

Query Match 35.2%; Score 32; DB 2; Length 96;

Best Local Similarity 50.0%; Pred. No. 2.5e+02;

Matches 6; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

QY 1 RRRVYDALNVLMMNI 16

DB 78 RNKLEAYNELMKQV 93

RESULT 15

G70401

flagellar hook-basal body protein - Aquifex aeolicus

C;Species: Aquifex aeolicus

C;Date: 08-May-1998 #sequence\_revision 08-May-1998 #text\_change 17-Mar-2000

C;Accession: G70401

R;Decker, G.; Warren, P.V.; Gaasterland, T.; Young, W.G.; Lenox, A.L.; Graham, D.E.; O.V. Nature 392, 353-358, 1998

A;Title: The complete genome of the hyperthermophilic bacterium Aquifex aeolicus.

A;Reference number: A70300; MUID:98196666; PMID:9537320

A;Accession: G70401

QY 3 RRVYDALNVLMMNIIS 18

DB 2 RIVAFEMSMEKDIIS 17

RESULT 13

H84088

hypothetical protein BH3512 [imported] - Bacillus halodurans (strain C-125)

C;Species: Bacillus halodurans

C;Date: 01-Dec-2000 #sequence\_revision 01-Dec-2000 #text\_change 15-Jun-2001

C;Accession: H84088

R;Takami, H.; Nakasone, K.; Takaki, Y.; Maeno, G.; Sasaki, R.; Masui, N.; Fujii, F.; Hira, Nucleic Acids Res. 28, 4317-4331, 2000

A;Title: Complete genome sequence of the alkaliphilic bacterium Bacillus halodurans and

A;Reference number: A83650; MUID:20512582; PMID:11058132

A;Accession: H84088

A;Status: preliminary

A;Molecule type: DNA

A;Residues: 1-83 <STO>

A;Cross-references: GB:AP001519; GB:BA000004; NID:gl0176109; PIDN:BA07231.1; GSPDB:GN00

A;Experimental source: strain C-125

C;Genetics:

A;Gene: BH3512

Query Match 35.2%; Score 32; DB 2; Length 83;

Best Local Similarity 60.0%; Pred. No. 2.1e+02;

Matches 6; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 2 RRVYDALNVL 11

DB 14 REIYDTLNLG 23

RESULT 14

S63609

extracellular secretory protein nucE - Serratia marcescens

C;Species: Serratia marcescens

C;Date: 28-Oct-1996 #sequence\_revision 13-Mar-1997 #text\_change 08-Oct-1999

C;Accession: S63609

R;Jin, S.; Chen, Y.; Christie, G.E.; Benedik, M.J. J. Mol. Biol. 256, 264-278, 1996

A;Title: Regulation of the Serratia marcescens extracellular nuclease: positive control

A;Reference number: S63609; MUID:96174474; PMID:8594195

A;Accession: S63609

A;Status: preliminary

A;Molecule type: DNA

A;Residues: 1-89 <JIN>

A;Cross-references: EMBL:U11698; NID:9509550; PIDN:AAA98439.1; PID:9509551

C;Genetics:

A;Gene: nucE

Query Match 35.2%; Score 32; DB 2; Length 89;

Best Local Similarity 55.6%; Pred. No. 2.3e+02;

Matches 5; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 1 RRRVYDALN 9

DB 80 RKKIYDECN 88

RESULT 15

G70401

flagellar hook-basal body protein - Aquifex aeolicus

C;Species: Aquifex aeolicus

C;Date: 08-May-1998 #sequence\_revision 08-May-1998 #text\_change 17-Mar-2000

C;Accession: G70401

R;Decker, G.; Warren, P.V.; Gaasterland, T.; Young, W.G.; Lenox, A.L.; Graham, D.E.; O.V. Nature 392, 353-358, 1998

A;Title: The complete genome of the hyperthermophilic bacterium Aquifex aeolicus.

A;Reference number: A70300; MUID:98196666; PMID:9537320

A;Accession: G70401

QY 8 LNVLMAMNIISK 19  
| | | | |  
Db 60 LKALMQLNIVEK 71

## RESULT 18

S01102

hypothetical protein 1 - fruit fly (*Drosophila melanogaster*)C;Species: *Drosophila melanogaster*

C;Date: 30-Jun-1989 #sequence\_revision 30-Jun-1989 #text\_change 21-Jul-2000

C;Accession: S01102

R;Eveleth, D.D.; Marsh, J.L.

Mol. Gen. Genet. 209, 290-298, 1987

A;Title: Overlapping transcription units in *Drosophila*: sequence and structure of the C8

A;Reference number: S01102; MUID:88038375; PMID:3478553

A;Accession: S01102

A;Molecule type: DNA

A;Residues: 1-40 &lt;EVE&gt;

A;Cross-references: EMBL:X05991; NID:g7759; PIDN:CAA29405.1; PID:g2598388

C;Genetics:

A;Gene: C8

A;Cross-references: FlyBase:FBgn0002036

A;Introns: 3/2

Query Match 34.6%; Score 31.5; DB 2; Length 40;

Best Local Similarity 52.6%; Pred. No. 1.2e+02;

Matches 10; Conservative 2; Mismatches 6; Indels 1; Gaps 1;

QY 1 RRRVYDALNVLMAMNIISK 19

| | | | |

Db 9 RRSLYNA-RVLOADNIGDK 26

## RESULT 19

D83924

hypothetical protein BH2196 [imported] - *Bacillus halodurans* (strain C-125)C;Species: *Bacillus halodurans*

C;Date: 01-Dec-2000 #sequence\_revision 01-Dec-2000 #text\_change 15-Jun-2001

C;Accession: D83924

R;Takami, H.; Nakasone, K.; Takaki, Y.; Maeno, G.; Sasaki, R.; Masui, N.; Fuji, F.; Hira

Nucleic Acids Res. 28, 4317-4331, 2000

A;Title: Complete genome sequence of the alkaliphilic bacterium *Bacillus halodurans* and

A;Reference number: A83650; MUID:20512582; PMID:11058132

A;Accession: D83924

A;Status: preliminary

A;Molecule type: DNA

A;Residues: 1-35 &lt;STO&gt;

A;Cross-references: GB:AP001514; GB:BA0000004; NID:gl0174613; PIDN:BA05915.1; GSPDB:GN00

A;Experimental source: strain C-125

C;Genetics:

A;Gene: BH2196

Query Match 34.1%; Score 31; DB 2; Length 35;

Best Local Similarity 35.7%; Pred. No. 1.2e+02;

Matches 5; Conservative 6; Mismatches 3; Indels 0; Gaps 0;

QY 4 VYDALNVLMAMNII 17

| | | | |

Db 3 VYETLSIVVQNSI 16

## RESULT 20

A42766

jararatribrase II (EC 3.4.24.-) - jararaca (fragment)

C;Species: Bothrops jararaca (jararaca)

C;Date: 04-Mar-1993 #sequence\_revision 18-Nov-1994 #text\_change 25-Mar-1998

C;Accession: A42766

R;Maruyama, M.; Sugiki, M.; Yoshida, E.; Mihara, H.; Nakajima, N.

Toxicol 30, 853-864, 1992

A;Title: Purification and characterization of two fibrinolytic enzymes from *Bothrops jar*

A;Reference number: A42766; MUID:92397346; PMID:1523677

A;Accession: A42766

A;Status: preliminary

A;Molecule type: protein  
A;Residues: 1-49 <MAR>  
A;Experimental source: venom  
A;Note: sequence extracted from NCBI backbone (NCBIP:113117)  
C;Superfamily: atrolysin C  
C;Keywords: hydrolase; metalloproteinase; zinc

Query Match 34.1%; Score 31; DB 2; Length 49;  
Best Local Similarity 36.4%; Pred. No. 1.8e+02;  
Matches 4; Conservative 5; Mismatches 2; Indels 0; Gaps 0;

QY 1 RRRVYDALNVL 11

| | | | |

Db 30 RRRHQVMNIM 40

## RESULT 21

R5EC32

ribosomal protein L32 [validated] - *Escherichia coli* (strain K-12)C;Species: *Escherichia coli*

C;Date: 24-Apr-1984 #sequence\_revision 30-Jun-1991 #text\_change 01-Mar-2002

C;Accession: JV0048; A02832; A30419; F64852

R;Tanaka, Y.; Tsujimura, A.; Fujita, N.; Isono, S.; Isono, K.

J. Bacteriol. 171, 5707-5712, 1989

A;Title: Cloning and analysis of an *Escherichia coli* operon containing the rpmF gene for

A;Reference number: JV0048; MUID:90008815; PMID:2477362

A;Accession: JV0048

A;Molecule type: DNA

A;Residues: 1-57 &lt;TAN&gt;

A;Cross-references: GB:M29698; NID:gl47710; PIDN:AAA24575.1; PID:gl47712

A;Experimental source: strain K-12

R;Wittmann-Liebold, B.; Greuer, B.; Pannenbecker, R.

Hoppe-Seyler's Z. Physiol. Chem. 356, 1977-1979, 1975

A;Title: The primary structure of protein L32 from the 50S subunit of *Escherichia coli* r

A;Reference number: A02832; MUID:76119562; PMID:765258

A;Accession: A02832

A;Molecule type: protein

A;Residues: 2-57 &lt;WIT&gt;

A;Experimental source: strain K

R;Vinokurov, L.M.; Alakhov, Y.B.; Golov, E.A.; Ovchinnikov, Y.A.

Bioorg. Khim. 2, 1013-1017, 1976

A;Reference number: A30419

A;Accession: A30419

A;Molecule type: protein

A;Residues: 2-57 &lt;VIN&gt;

A;Experimental source: strain MR8-600

R;Blattner, F.R.; Plunkett III, G.; Bloch, C.A.; Perna, N.T.; Burland, V.; Riley, M.; Col

A.; Rose, D.J.; Mau, B.; Shao, Y.

Science 277, 1453-1462, 1997

A;Title: The complete genome sequence of *Escherichia coli* K-12.

A;Reference number: A64720; MUID:97426617; PMID:9278503

A;Accession: F64852

A;Status: nucleic acid sequence not shown; translation not shown

A;Molecule type: DNA

A;Residues: 1-57 &lt;BLAT&gt;

A;Cross-references: GB:AE000209; GB:U00096; NID:gl787322; PIDN:AAC74173.1; PID:gl787330;

A;Experimental source: strain K-12, substrain MG1655

R;Arnold, R.J.; Reilly, J.R.

Anal. Biochem. 269, 105-112, 1999

A;Title: Observation of *Escherichia coli* ribosomal proteins and their posttranslational n

A;Reference number: A59071; MUID:99196679; PMID:10094780

A;Contents: annotation; mass spectrographic analysis

A;Note: mass spectrographic analysis of post-translational modifications; any acid labile

C;Genetics:

A;Gene: rpmf

A;Map position: 24 min

C;Complex: the ribosome is composed of the large (50S) and small (30S) subunit; the large

S rRNA and 22 distinct proteins

C;Complex: large subunit ribosomal proteins: L1 (PIR:R5EC1), L3 (PIR:R5EC3), L2 (PIR:R5EC

(PIR:R5EC11), L13 (PIR:R5EC13), L14 (PIR:R5EC14), L15 (PIR:R5EC15), L16 (PIR:R5EC16), L1

C;Function:

A;Pathway: protein biosynthesis

C;Superfamily: *Escherichia coli* ribosomal protein L32

C;Keywords: protein biosynthesis; ribosome  
F;2-57/Product: ribosomal protein L32 #status experimental <MAT>

Query Match 34.1%; Score 31; DB 1; Length 57;  
Best Local Similarity 33.3%; Pred. No. 2.1e+02;  
Matches 5; Conservative 6; Mismatches 4; Indels 0; Gaps 0;

QY 2 RRVYDALNVLMANNI 16  
||:||||:||||:  
Db 16 RRSHDALTAVTSLSV 30

## RESULT 22

C90812  
50S ribosomal subunit protein L32 [imported] - Escherichia coli (strain O157:H7, substra

C;Species: Escherichia coli  
C;Date: 18-Jul-2001 #sequence\_revision 18-Jul-2001 #text\_change 17-May-2002

C;Accession: C90812  
R;Hayashi, T.; Makino, K.; Ohnishi, M.; Kurokawa, K.; Ishii, K.; Yokoyama, K.; Han, C.G.

gasawara, N.; Yasunaga, T.; Kuwara, S.; Shiba, T.; Hattori, M.; Shinagawa, H.  
DNA Res. 8, 11-22, 2001

A;Title: Complete genome sequence of enterohemorrhagic Escherichia coli O157:H7 and gene

A;Reference number: A99629; MUID:21156231; PMID:11258796

A;Accession: C90812

A;Status: preliminary

A;Molecule type: DNA

A;Residues: 1-57 <HAY>

A;Cross-references: GB:BA000007; PIDN:BA034890.1; PID:gl3360931; GSPDB:GN00154

A;Experimental source: strain O157:H7, substrain RIMD 0509952

C;Genetics:

A;Gene: ECs1467

C;Superfamily: Escherichia coli ribosomal protein L32

Query Match 34.1%; Score 31; DB 2; Length 57;  
Best Local Similarity 33.3%; Pred. No. 2.1e+02;  
Matches 5; Conservative 6; Mismatches 4; Indels 0; Gaps 0;

QY 2 RRVYDALNVLMANNI 16  
||:||||:||||:  
Db 16 RRSHDALTAVTSLSV 30

## RESULT 23

G85671  
50S ribosomal subunit protein L32 [imported] - Escherichia coli (strain O157:H7, substra

C;Species: Escherichia coli

C;Date: 16-Feb-2001 #sequence\_revision 16-Feb-2001 #text\_change 17-May-2002

C;Accession: G85671

R;Perna, N.T.; Plunkett III, G.; Burland, V.; Mau, B.; Glasner, J.D.; Rose, D.J.; Mayhew

iller, L.; Grotbeck, E.J.; Davis, N.W.; Lim, A.; Dimalanta, E.; Potamouisis, K.; Apodaca,

Nature 409, 529-533, 2001

A;Title: Genome sequence of enterohemorrhagic Escherichia coli O157:H7.

A;Reference number: A85480; MUID:21074935; PMID:11206551

A;Accession: G85671

A;Status: preliminary

A;Molecule type: DNA

A;Residues: 1-57 <STO>

A;Cross-references: GB:AE005174; NID:gl2514633; PIDN:AG55835.1; GSPDB:GN00145; UWGP:Z17

A;Experimental source: strain O157:H7, substrain EDL933

C;Genetics:

A;Gene: rpmF

C;Superfamily: Escherichia coli ribosomal protein L32

Query Match 34.1%; Score 31; DB 2; Length 57;  
Best Local Similarity 33.3%; Pred. No. 2.1e+02;  
Matches 5; Conservative 6; Mismatches 4; Indels 0; Gaps 0;

QY 2 RRVYDALNVLMANNI 16  
||:||||:||||:  
Db 16 RRSHDALTAVTSLSV 30

## RESULT 24

## AH0641

50S ribosomal protein L32 [imported] - Salmonella enterica subsp. enterica serovar Typhi

C;Species: Salmonella enterica subsp. enterica serovar Typhi

A;Note: this species has also been called Salmonella typhi

C;Date: 09-Nov-2001 #sequence\_revision 09-Nov-2001 #text\_change 18-Nov-2002

C;Accession: AH0641

R;Parkhill, J.; Dougan, G.; James, K.D.; Thomson, N.R.; Pickard, D.; Wain, J.; Churcher,

th, T.; Connor, P.; Cronin, A.; Davis, P.; Davies, R.M.; Dowd, L.; White, N.; Farrar,

S.; Moule, S.; O'Gaora, P.

Nature 413, 848-852, 2001

A;Authors: Parry, C.; Quail, M.; Rutherford, K.; Simmonds, M.; Skelton, J.; Stevens, K.;

A;Title: Complete genome sequence of a multiple drug resistant Salmonella enterica serovi

A;Reference number: AB0502; MUID:21534947; PMID:11677608

A;Accession: AH0641

A;Status: preliminary

A;Molecule type: DNA

A;Residues: 1-57 <PAR>

A;Cross-references: GB:AU513382; PIDN:CAD08315.1; PID:gl6502361; GSPDB:GN00176

C;Genetics:

A;Gene: STY1230

C;Superfamily: Escherichia coli ribosomal protein L32

Query Match 34.1%; Score 31; DB 2; Length 57;  
Best Local Similarity 33.3%; Pred. No. 2.1e+02;  
Matches 5; Conservative 6; Mismatches 4; Indels 0; Gaps 0;

QY 2 RRVYDALNVLMANNI 16  
||:||||:||||:  
Db 16 RRSHDALTAVTSLSV 30

## RESULT 25

## AI2346

hypothetical protein asl4328 [imported] - Nostoc sp. (strain PCC 7120)

C;Species: Nostoc sp. PCC 7120

A;Note: Nostoc sp. strain PCC 7120 is a synonym of Anabaena sp. strain PCC 7120

C;Date: 14-Dec-2001 #sequence\_revision 14-Dec-2001 #text\_change 09-Dec-2002

C;Accession: AI2346

R;Kaneko, T.; Nakamura, Y.; Wolk, C.P.; Kuritz, T.; Sasamoto, S.; Watanabe, A.; Iriguchi,

Nakazaki, N.; Shimpo, S.; Sugimoto, M.; Takazawa, M.; Yamada, M.; Yasuda, M.; Tabata, S.

DNA Res. 8, 205-213, 2001

A;Title: Complete Genomic Sequence of the Filamentous Nitrogen-fixing Cyanobacterium Anati

A;Reference number: AB1807; MUID:21595285; PMID:11759840

A;Accession: AI2346

A;Status: preliminary

A;Molecule type: DNA

A;Residues: 1-68 <KUR>

A;Cross-references: GB:BA000019; PIDN:BA076027.1; PID:gl7133464; GSPDB:GN00179

A;Experimental source: strain PCC 7120

C;Genetics:

A;Gene: asl4328

Query Match 34.1%; Score 31; DB 2; Length 68;  
Best Local Similarity 66.7%; Pred. No. 2.6e+02;  
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 5 YDALNVLMA 13  
|:|:|:|:  
Db 43 YEAELELMA 51

## RESULT 26

## I83374

protein-tyrosine kinase - mouse (fragment)

C;Species: Mus musculus (house mouse)

C;Date: 29-May-1998 #sequence\_revision 29-May-1998 #text\_change 05-Nov-1999

C;Accession: I83374

R;Hebert, B.; Bergeron, J.; Tijssen, P.; Potworowski, E.F.

Gene 143, 257-260, 1994

A;Title: Protein tyrosine kinases transcribed in a murine thymic medullary epithelial ce

A;Reference number: I60294; MUID:94266162; PMID:8206383

A;Accession: I83374

A;Status: preliminary; translated from GB/EMBL/DBJ

A;Molecule type: mRNA  
A;Residues: 1-70 <RES>  
A;Cross-references: GB:L25762; NID:G413748; PIDN:AAA40020.1; PID:G554265  
C;Genetics:  
A;Gene: PTK  
C;Superfamily: protein kinase homology (fragment) <KIN>  
F;1-70/Domain: protein kinase homology

Query Match 34.1%; Score 31; DB 2; Length 70;  
Best Local Similarity 46.2%; Pred. No. 2.7e+02;  
Matches 6; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

QY 7 ALNVLAMNIIISK 19  
|||:|:|:  
Db 9 AANILVGENLICK 21  
|||:|:|:

RESULT 27  
T10345  
hypothetical protein 76 - Orgyia pseudotsugata nuclear polyhedrosis virus  
C;Species: Orgyia pseudotsugata nuclear polyhedrosis virus, OpMNVP  
C;Date: 16-Jul-1999 #sequence\_revision 16-Jul-1999 #text\_change 11-May-2000  
C;Accession: T10345  
R;Ahrens, C.A.; Russell, R.R.; Funk, C.J.; Evans, J.; Harwood, S.; Rohrmann, G.F.  
Virology 229, 381-399, 1997  
A;Title: The sequence of the Orgyia pseudotsugata multinucleocapsid nuclear polyhedrosis virus  
A;Reference number: Z17011; MUID:97271300; PMID:9126251  
A;Accession: T10345  
A;Status: preliminary; translated from GB/EMBL/DBJ  
A;Molecule type: DNA  
A;Residues: 1-84 <AHR>  
A;Cross-references: EMBL:U75930; NID:G2934903; PID:G1911322

Query Match 34.1%; Score 31; DB 2; Length 84;  
Best Local Similarity 31.6%; Pred. No. 3.2e+02;  
Matches 6; Conservative 7; Mismatches 4; Indels 2; Gaps 1;

QY 1 RRVYDA--LNLVAMNII 17  
|||:|:|:|:  
Db 56 RRSFYESLKLNCIVCNVL 74  
|||:|:|:|:

RESULT 28  
S74489  
hypothetical protein ssr1853 - Synechocystis sp. (strain PCC 6803)  
C;Species: Synechocystis sp.  
A;Variety: PCC 6803  
C;Date: 25-Apr-1997 #sequence\_revision 25-Apr-1997 #text\_change 08-Oct-1999  
C;Accession: S74489  
R;Kaneko, T.; Sato, S.; Kotani, H.; Tanaka, A.; Asamizu, Y.; Miyajima, N.;  
O. K.; Okumura, S.; Shimpo, S.; Takeuchi, C.; Wada, T.; Watanabe, A.; Yamada, M.; Yasuda  
DNA Res. 3, 109-136, 1996  
A;Title: Sequence analysis of the genome of the unicellular cyanobacterium Synechocystis  
s.  
A;Reference number: S74322; MUID:97061201; PMID:8905231  
A;Accession: S74489  
A;Status: nucleic acid sequence not shown; translation not shown  
A;Molecule type: DNA  
A;Residues: 1-95 <KAN>  
A;Cross-references: EMBL:D90899; GB:AB001339; NID:G1651650; PIDN:BAAL6641.1; PID:d101737  
A;Note: the nucleotide sequence was submitted to the EMBL Data Library, June 1996

Query Match 34.1%; Score 31; DB 2; Length 95;  
Best Local Similarity 54.5%; Pred. No. 3.7e+02;  
Matches 6; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 1 RRVYDALNVL 11  
|||:|:|:  
Db 38 RQRIQVALNAL 48  
|||:|:|:

RESULT 29  
H89910

conserved hypothetical protein SAll186 [imported] - Staphylococcus aureus (strain N315)  
C;Species: Staphylococcus aureus  
C;Date: 10-May-2001 #sequence\_revision 10-May-2001 #text\_change 09-Dec-2002  
C;Accession: H89910  
R;Kuroda, M.; Ohta, T.; Uchiyama, I.; Baba, T.; Yuzawa, H.; Kobayashi, I.; Cui, L.; Oguc  
ma, A.; Mizutani-Ui, Y.; Kobayashi, N.; Sawano, T.; Inoue, R.; Kaico, C.; Sekimizu, K.;  
C.; Shiba, T.; Hattori, M.; Ogasawara, N.; Hayashi, H.; Hiramatsu, K.  
Lancet 357, 1225-1240, 2001  
A;Title: Whole genome sequencing of methicillin-resistant Staphylococcus aureus.  
A;Reference number: A89758; MUID:21311952; PMID:11418146  
A;Accession: H89910  
A;Status: preliminary  
A;Molecule type: DNA  
A;Residues: 1-98 <KUR>  
A;Cross-references: GB:BA000018; PID:G13701149; PIDN:BAB42444.1; GSPDB:GN00149  
A;Experimental source: strain N315  
C;Genetics:  
A;Gene: SAll186  
C;Superfamily: uncharacterized conserved protein

Query Match 34.1%; Score 31; DB 2; Length 98;  
Best Local Similarity 66.7%; Pred. No. 3.8e+02;  
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 5 YDALNVLMMA 13  
|||:|:|:  
Db 62 YDDLNVVVA 70  
|||:|:|:

RESULT 30  
B75624  
transposase, pseudogene - Deinococcus radiodurans (strain R1)  
C;Species: Deinococcus radiodurans  
C;Date: 03-Dec-1999 #sequence\_revision 03-Dec-1999 #text\_change 31-Mar-2000  
C;Accession: B75624  
R;White, O.; Eisen, J.A.; Heidelberg, J.F.; Hickey, E.K.; Peterson, J.D.; Dodson, R.J.;  
M.; Shen, M.; Vamathevan, J.J.; Lam, P.; McDonald, L.; Utterback, T.; Zalewski, C.; Ma  
S.; Smith, H.O.; Venter, J.C.; Fraser, C.M.  
Science 286, 1571-1577, 1999  
A;Title: Genome sequence of the radioresistant bacterium Deinococcus radiodurans R1.  
A;Reference number: A75250; MUID:20036896; PMID:10567266  
A;Accession: B75624  
A;Status: preliminary  
A;Molecule type: DNA  
A;Residues: 1-66 <WHI>  
A;Cross-references: GB:AE001826; NID:G6460827; PIDN:AAF12601.1; PID:G6460897; TIGR:DRB00  
A;Experimental source: strain R1  
C;Genetics:  
A;Gene: DRB0055  
A;Map position: megaplasmid  
A;Genome: plasmid  
A;Note: plasmid MPI

Query Match 33.5%; Score 30.5; DB 2; Length 66;  
Best Local Similarity 44.4%; Pred. No. 3e+02;  
Matches 8; Conservative 3; Mismatches 4; Indels 3; Gaps 1;

QY 1 RRVYDALNVLAMNIIIS 18  
|||:|:|:|:  
Db 26 KRRAY---NALLAQVILS 40  
|||:|:|:|:

RESULT 31  
G82881  
acyl carrier protein UUS06 [imported] - Ureaplasma urealyticum  
C;Species: Ureaplasma urealyticum  
C;Date: 18-Aug-2000 #sequence\_revision 20-Aug-2000 #text\_change 20-Aug-2000  
C;Accession: G82881  
R;Glass, J.I.; Lefkowitz, E.J.; Glass, J.S.; Heiner, C.R.; Chen, E.Y.; Cassell, G.H.  
submitted to GenBank, February 2000  
A;Description: The complete sequence of Ureaplasma urealyticum: Alternate views of a mini  
A;Reference number: A82870  
A;Accession: G82881

A;Status: preliminary  
A;Molecule type: DNA  
A;Residues: 1-77 <GLA>  
A;Cross-references: GB:AE002149; GB:AF222894; PIDN:AAF30918.1; GSPDB:GN001  
A;Experimental source: serovar 3; biovar 1  
C;Genetics:  
A;Gene: acpP; UUS06  
A;Genetic code: SGC3

Query Match 33.5%; Score 30.5; DB 2; Length 77;  
Best Local Similarity 39.1%; Pred. No. 3.6e+02;  
Matches 9; Conservative 2; Mismatches 3; Indels 9; Gaps 1;

Qy 6 DALNV-----LMAMNLIISK 19  
| ||| | ||| | |  
Db 23 DNLNIELKSLGIDLSLNNLIMK 45

RESULT 32  
PA0105  
heat shock protein dnaJ type - fungus (Fusarium sporotrichioides) (fragment)  
C;Species: Fusarium sporotrichioides  
C;Date: 20-Feb-1995 #sequence\_revision 20-Feb-1995 #text\_change 23-Mar-2001  
C;Accession: PA0105  
R;Chow, L.P.; Fukaya, N.; Sugiura, Y.; Ueno, Y.; Tabuchi, K.; Taugita, A.  
submitted to JIPID, October 1994  
A;Description: Two dimensional polyacrylamide gel electrophoresis of Fusarium sporotrichioides  
A;Reference number: PA0051  
A;Accession: PA0105  
A;Molecule type: protein  
A;Residues: 1-15 <CHO>  
C;Keywords: heat shock; stress-induced protein

Query Match 33.0%; Score 30; DB 2; Length 15;  
Best Local Similarity 62.5%; Pred. No. 73;  
Matches 5; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Qy 3 RYVDALNV 10  
: || || |  
Db 5 KLYDTLNV 12

RESULT 33  
E89904  
hypothetical protein [imported] - Staphylococcus aureus (strain N315)  
C;Species: Staphylococcus aureus  
C;Date: 10-May-2001 #sequence\_revision 10-May-2001 #text\_change 22-Oct-2001  
C;Accession: E89904  
R;Kuroda, M.; Ohta, T.; Uchiyama, I.; Baba, T.; Yuzawa, H.; Kobayashi, I.; Cui, L.; Ogino, A.; Mizutani-Ui, Y.; Kobayashi, N.; Sawano, T.; Inoue, R.; Kaito, C.; Sekimizu, K.; C.; Shiba, T.; Hattori, M.; Ogasawara, N.; Hayashi, H.; Hiramatsu, K.  
Lancet 357, 1225-1240, 2001  
A;Title: Whole genome sequencing of methicillin-resistant Staphylococcus aureus.  
A;Reference number: R89758; MUID:21311952; PMID:11418146  
A;Accession: E89904  
A;Status: preliminary  
A;Molecule type: DNA  
A;Residues: 1-29 <KUR>  
A;Cross-references: GB:BA000018; PID:g13701098; PIDN:BA042393.1; GSPDB:GN00149  
A;Experimental source: strain N315  
C;Genetics:  
A;Gene: SAS038

Query Match 33.0%; Score 30; DB 2; Length 29;  
Best Local Similarity 55.6%; Pred. No. 1.5e+02;  
Matches 5; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

Qy 4 VYDNLNVLM 12  
: || || |  
Db 7 LYDTLNTLL 15

RESULT 34



A;Experimental source: strain 1021, megaplasmid pSymb  
 R;Galibert, F.; Finan, T.M.; Long, S.R.; Puhler, A.; Abola, P.; Ampe, P.; Barloy-Hubler, P.; Chain, P.; Cowie, A.; Davis, R.W.; Dreano, S.; Federspiel, N.A.; Fisher, R.F.; L.; Hyman, R.W.; Jones, T.  
 Science 293, 668-672, 2001  
 A;Authors: Kahn, D.; Kahn, M.L.; Kalman, S.; Keating, D.H.; Kiss, E.; Komp, C.; Lelaure, hebaule, P.; Vandenbol, M.; Vorholter, F.J.; Weidner, S.; Wells, D.H.; Wong, K.; Yen, K.  
 A;Title: The composite genome of the legume symbiont *Sinorhizobium meliloti*.  
 A;Reference number: A96039; MUID:21368234; PMID:11474104  
 A;Contents: annotation  
 C;Genetics:  
 A;Gene: SMB20061  
 A;Genome: plasmid

Query Match 33.0%; Score 30; DB 2; Length 51;  
 Best Local Similarity 58.3%; Pred. No. 2.9e+02;  
 Matches 7; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 2 RRYVDALNVLMMA 13  
 |||:|||||  
 Db 28 RLVAEALNLLFA 39

RESULT 37  
 D82422  
 hypothetical protein VCA0746 [imported] - *Vibrio cholerae* (strain N16961 serogroup O1)  
 C;Species: *Vibrio cholerae*  
 C;Date: 18-Aug-2000 #sequence\_revision 20-Aug-2000 #text\_change 02-Feb-2001  
 C;Accession: D82422  
 R;Heidelberg, J.F.; Eisen, J.A.; Nelson, W.C.; Clayton, R.A.; Gwinn, M.L.; Dodson, R.J.; chardson, D.; Ermolaeva, M.D.; Vamathevan, J.; Bass, S.; Qin, H.; Dragoi, I.; Sellers, R. l, R.R.; Mekalanos, J.J.; Venter, J.C.; Fraser, C.M.  
 Nature 406, 477-483, 2000  
 A;Title: DNA Sequence of both chromosomes of the cholera pathogen *Vibrio cholerae*.  
 A;Reference number: A82035; MUID:20406833; PMID:10952301  
 A;Accession: D82422  
 A;Status: preliminary  
 A;Molecule type: DNA  
 A;Residues: 1-53 <HEI>  
 A;Cross-references: GB:AE004403; GB:AE003853; NID:99658159; PIDN:AAF96644.1; GSPDB:GN001  
 A;Experimental source: serogroup O1; strain N16961; biotype El Tor  
 C;Genetics:  
 A;Gene: VCA0746  
 A;Map position: 2

Query Match 33.0%; Score 30; DB 2; Length 53;  
 Best Local Similarity 41.7%; Pred. No. 2.9e+02;  
 Matches 5; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

QY 2 RRYVDALNVLMMA 13  
 |||:|||||  
 Db 7 RNMVDTFNVRLS 18

RESULT 38  
 AC0657  
 hypothetical protein STY1359 [imported] - *Salmonella enterica* subsp. *enterica* serovar Ty  
 C;Species: *Salmonella enterica* subsp. *enterica* serovar Typhi  
 A;Note: this species has also been called *Salmonella typhi*  
 C;Date: 09-Nov-2001 #sequence\_revision 09-Nov-2001 #text\_change 18-Nov-2002  
 C;Accession: AC0657  
 R;Parkhill, J.; Dougan, G.; James, K.D.; Thomson, N.R.; Pickard, D.; Wain, J.; Churcher, th, T.; Connerton, P.; Cronin, A.; Davis, P.; Davies, R.M.; Dowd, L.; White, N.; Farrar, . S.; Moule, S.; O'Gaora, P.  
 Nature 413, 848-852, 2001  
 A;Authors: Parry, C.; Quail, M.; Rutherford, K.; Simmonds, M.; Skelton, J.; Stevens, K. A;Title: Complete genome sequence of a multiple drug resistant *Salmonella enterica* serov A;Reference number: AB0502; MUID:21534947; PMID:11677608  
 A;Accession: AC0657  
 A;Status: preliminary  
 A;Molecule type: DNA  
 A;Residues: 1-57 <PAR>  
 A;Cross-references: GB:AL513382; PIDN:CAD01628.1; PID:gl6502482; GSPDB:GN00176

C;Genetics:  
 A;Gene: STY1359

Query Match 33.0%; Score 30; DB 2; Length 57;  
 Best Local Similarity 50.0%; Pred. No. 3.2e+02;  
 Matches 6; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

QY 2 RRYVDALNVLMMA 13  
 |||:|||||  
 Db 23 RRFYKVCNIIIA 34

RESULT 39  
 D83797  
 hypothetical protein BHI180 [imported] - *Bacillus halodurans* (strain C-125)  
 C;Species: *Bacillus halodurans*  
 C;Date: 01-Dec-2000 #sequence\_revision 01-Dec-2000 #text\_change 15-Jun-2001  
 C;Accession: D83797  
 R;Takami, H.; Nakasone, K.; Takaki, Y.; Maeno, G.; Sasaki, R.; Masui, N.; Fujii, F.; Hiran Nucleic Acids Res. 28, 4317-4331, 2000  
 A;Title: Complete genome sequence of the alkaliphilic bacterium *Bacillus halodurans* and A;Reference number: A83650; MUID:20512582; PMID:11058132  
 A;Accession: D83797  
 A;Status: preliminary  
 A;Molecule type: DNA  
 A;Residues: 1-59 <STO>  
 A;Cross-references: GB:AP001511; GB:BA000004; NID:gl0173727; PIDN:BA04899.1; GSPDB:GN001  
 A;Experimental source: strain C-125  
 C;Genetics:  
 A;Gene: BHI180

Query Match 33.0%; Score 30; DB 2; Length 59;  
 Best Local Similarity 31.2%; Pred. No. 3.3e+02;  
 Matches 5; Conservative 6; Mismatches 5; Indels 0; Gaps 0;

QY 3 RRYVDALNVLMMAIIS 18  
 |||:|||||  
 Db 36 KVFRTNIIIMTILLTS 51

RESULT 40  
 A84215  
 hypothetical protein Vng0568c [imported] - *Halobacterium* sp. NRC-1  
 C;Species: *Halobacterium* sp. NRC-1  
 C;Date: 02-Feb-2001 #sequence\_revision 02-Feb-2001 #text\_change 02-Feb-2001  
 C;Accession: A84215  
 R;Ng, W.V.; Kennedy, S.P.; Mahairas, G.G.; Berquist, B.; Pan, M.; Shukla, H.D.; Lasky, S. ; Leithausner, B.; Keller, K.; Cruz, R.; Danson, M.J.; Hough, D.W.; Maddocks, D.G.; Jablor Jung, K.H.; Alam, M.; Freitas, T.  
 Proc. Natl. Acad. Sci. U.S.A. 97, 12176-12181, 2000  
 A;Authors: Hou, S.; Daniels, C.J.; Dennis, P.P.; Omer, A.D.; Ebhardt, H.; Lowe, T.M.; Li A;Title: Genome sequence of *Halobacterium* species NRC-1.  
 A;Reference number: A84160; MUID:20504483; PMID:11016950  
 A;Accession: A84215  
 A;Status: preliminary  
 A;Molecule type: DNA  
 A;Residues: 1-60 <STO>  
 A;Cross-references: GB:AE004437; NID:gl0580165; PIDN:AAG19085.1; GSPDB:GN00138  
 C;Genetics:  
 A;Gene: VNG0568C

Query Match 33.0%; Score 30; DB 2; Length 60;  
 Best Local Similarity 53.8%; Pred. No. 3.3e+02;  
 Matches 7; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

QY 4 VYDALNVLMAMNI 16  
 |||:|||||  
 Db 37 VYALLNPLMSIAI 49

RESULT 41  
 H81811  
 hypothetical protein NMA1853 [imported] - *Neisseria meningitidis* (strain Z2491 serogroup

C/Species: Molluscum contagiosum virus 1  
C/Date: 05-Nov-1999 #sequence\_revision 05-Nov-1999 #text\_change 11-May-2000  
C/Accession: T30649  
R/Senkevitch, T.G.; Bugert, J.J.; Sisler, J.R.; Koonin, E.V.; Darai, G.; Moss, B.  
Science 273, 813-816, 1996  
A/Title: Genome sequence of a human tumorigenic poxvirus: Prediction of specific host re  
A/Reference number: Z20876; MUID:96325459; PMID:8670425  
A/Accession: T30649  
A/Status: preliminary; translated from GB/EMBL/DBJ  
A/Molecule type: DNA  
A/Residues: 1-82 <SEN>  
A/Cross-references: EMBL:U60315; PIDN:AAC55175.1  
C/Genetics:  
A/Note: MC047L

Query Match	33.0%;	Score 30;	DB 2;	Length 82;
Best Local Similarity	50.0%;	Pred. No. 4.7e+02;		
Matches	6;	Conservative	3;	Mismatches 3; Indels 0; Gaps 0;
Qy	7	ALNVLMMNNIIS	18	
		: :		
Dh	13	ALTVLMLLMVVS	24	

**rega protein - Serratia marcescens**

C:Species: *Serratia marcescens*  
C:Date: 31-Dec-1996 #sequence\_revision 31-Dec-1996 #text\_change 08-Oct-1999  
R:Accession: JC6046  
R:Ferret, S.; Viejo, M.B.; Guasch, J.F.; Enfedaque, J.; Regue, M.  
J. Bacteriol. 178, 951-960, 1996  
A:Title: Genetic evidence for an activator required for induction of colicin-like bacteriocin production in *Serratia marcescens*  
A:Reference number: JC6046; MUID:96165247; PMID:8576068  
A:Accession: JC6046  
A:Molecule type: DNA  
A:Residues: 1-88 <Ref>  
A:Cross-references: GB:U01763; NID:9965068; PIDN:AAA96011.1; PID:g965069  
C:Comment: This protein has two transmembrane helical domains and the charge-rich amino-terminal domain  
C:Genetics:  
A:Gene: regA  
C:Keywords: transmembrane protein

```

C:keywords: transmembrane protein

Query Match      33.0%;   Score 30;   DB 2;   Length 88;
Best Local Similarity 55.6%;   Pred. NO. 5.1e+02;
Matches 5;   Conservative 1;   Mismatches 3;   Indels 0;   Gaps 0;

QY      1 RRRVYDALN 9
      ||:||||
DB      79 RRIYDECN 87

RESULT 45
A38167
PhiH1 repressor - Halobacterium salinarum phage phi-H
C:Species: Halobacterium salinarum phage phi-H
C:Date: 28-Aug-1992 #sequence_revision 18-Sep-1992 #text_change 18-Jul-2001
C:Accession: A38167; S11641
Riken, R.; Hackett, N.R. 1991

```

J. Bacteriol. 173, 953-960, 1991  
 A>Title: Halobacterium halobium strains lysogenic for phage phiH contain a protein repressing the prophage  
 A:Reference number: A38167; PMID:91123222; PMID:1991733  
 A:Accession: A38167  
 A>Status: preliminary  
 A:Molecule type: DNA  
 A:Residues: 1-89 <KEN>  
 A:Cross-references: GB:X52504; NID:914814; PIDN:CAA36746.1; PID:914816  
 A>Note: the source is designated as Halobacterium halobium phase phiH1  
 A>Note: the authors translated the codon GCG for residue 14 as Arg, and CAA for residue 15 as Gln  
 C:Keywords: DNA binding; transcription regulation

C;Keywords: DNA binding; transcription regulation  
Query Match 33.0%; Score 30; DB 2; Length 89;  
Best Local Similarity 35.3%; Pred. NO. 5.1e+02;

Matches 6; Conservative 4; Mismatches 7; Indels 0; Gaps 0;

QY 1 RREYDANVLMANNII 17  
 |||:| |:  
 Db 40 RREISDRCTVLVDGRL 56

Search completed: February 17, 2004, 10:58:15  
 Job time : 12.4752 secs

**THIS PAGE BLANK (USPTO)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: February 17, 2004, 10:50:13 ; Search time 6.0198 Seconds  
(without alignments)  
148.428 Million cell updates/sec

Title: US-09-900-147-3

Perfect score: 91

Sequence: 1 RRRVVDALNVLMMNIISK 19

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 127863 seqs, 47026705 residues

Total number of hits satisfying chosen parameters: 13973

Minimum DB seq length: 0

Maximum DB seq length: 100

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt\_41.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	ID	Description
1	40	44.0	38	1 E2F1_RAT	O09139 rattus norv
2	32	35.2	70	1 YORA_TTV1	P19285 thermoprote
3	32	35.2	93	1 FLIE_AQUAE	P07242 aquifex aeo
4	31	34.1	56	1 RL32_ECOLI	P02435 escherichia
5	31	34.1	79	1 PLP_CRYNE	P82476 cryptococcu
6	31	34.1	84	1 Y073_NPVOP	O10326 orgyia pseu
7	30.5	33.5	77	1 ACPH_UREPA	Q9PPY4 ureaplasma
8	30	33.0	48	1 ATP8_CANPA	P17345 candida par
9	30	33.0	56	1 SCP2_MESMA	Q9NJF7 mesobuthus
10	30	33.0	56	1 SCP3_MESMA	Q9U8D1 mesobuthus
11	30	33.0	89	1 VREP_BPPHH	P22562 bacterioph
12	30	33.0	96	1 YFRC_PROVU	P20927 proteus vul
13	29.5	32.4	54	1 APR_HUMAN	Q13794 homo sapien
14	29.5	32.4	56	1 RL32_MYCGE	P47603 mycoplasma
15	29.5	32.4	56	1 RL32_MYCPN	P75238 mycoplasma
16	29	31.9	83	1 PSBE_NEPOL	Q9TKY1 nephroselmi
17	29	31.9	97	1 H41_BLEJA	P80737 blepharisma
18	29	31.9	98	1 NULM_DASNO	O21333 dasypus nov
19	28	30.8	34	1 COXG_THUOB	P80976 thunnus obe
20	28	30.8	42	1 RET5_BOVIN	P82708 bos taurus
21	28	30.8	61	1 RL32_MYCPU	Q98QN7 mycoplasma
22	28	30.8	67	1 YIFL_ECOLI	P73166 escherichia
23	28	30.8	77	1 REP1_ECOLI	P10396 escherichia
24	28	30.8	79	1 VI05_VACCC	P20500 vaccinia vi
25	28	30.8	79	1 VI05_VACCV	P12924 vaccinia vi
26	28	30.8	79	1 VI05_VARV	P33001 variola vir
27	28	30.8	91	1 VAPD_HAEIN	P71351 haemophilus
28	27.5	30.2	87	1 RT21_HUMAN	P82921 homo sapien
29	27.5	30.2	87	1 RT21_MOUSE	P58059 mus musculu
30	27.5	30.2	91	1 RS15_RICCN	Q92HV6 rickettsia
31	27	29.7	15	1 UC27_MAZE	P80633 zea mays (m
32	27	29.7	58	1 YC18_PORPU	P51366 porphyra pu
33	27	29.7	59	1 SECE_TREPA	O83263 treponema p

ALIGNMENTS

RESULT 1

ID	E2F1_RAT	STANDARD;	PRT;	38 AA.
AC	O09139;			
DT	01-NOV-1997 (Rel. 35, Created)			
DT	01-NOV-1997 (Rel. 35, Last sequence update)			
DT	15-JUL-1998 (Rel. 36, Last annotation update)			
DE	Transcription factor E2F1 (E2F-1) (Fragment).			
GN	E2F1.			
OS	Rattus norvegicus (Rat).			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Rattus.			
OX	NCBI_TaxID=10116;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RX	MEDLINE=96113578; PubMed=8673024;			
RA	Hosokawa Y., Yang M., Kaneko S., Tanaka M., Nakashima K.;			
RT	"Synergistic gene expressions of cyclin E, cdk2, cdk5 and E2F-1 during the prolactin-induced G1/S transition in rat Nb2 pre-T lymphoma cells."			
RT	Lymphoma cells."			
RL	Biochem. Mol. Biol. Int. 37:393-399(1995).			
CC	-I- FUNCTION: TRANSCRIPTIONAL ACTIVATOR THAT BINDS TO E2F SITES, THESE SITES ARE PRESENT IN THE PROMOTER OF MANY GENES WHOSE PRODUCTS ARE INVOLVED IN CELL PROLIFERATION MAY MEDIATE GROWTH FACTOR-INITIATED SIGNAL TRANSDUCTION (BY SIMILARITY).			
CC	-I- SUBUNIT: COMPONENT OF THE DRTF1/E2F TRANSCRIPTION FACTOR COMPLEX. BINDS COOPERATIVELY WITH DP-1 TO E2F SITES. INTERACTS PREPARENTIALLY WITH RETINOBLASTOMA PROTEIN RB1 THAT INHIBIT THE E2F TRANSACTION DOMAIN. SOME INTERACTION HAS ALSO BEEN FOUND WITH RETINOBLASTOMA RELATED PROTEIN 107 (BY SIMILARITY).			
CC	-I- SUBCELLULAR LOCATION: Nuclear.			
CC	-I- SIMILARITY: BELONGS TO THE E2F/DP FAMILY.			

-----  
This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
-----

EMBL; D63165; BAA09641.1; --	
HSSP; Q16254; 1CF7.	
TRANSFAC; T02952; --	
InterPro; IPR003316; E2F_TDP.	
Pfam; PF02319; E2F_TDP; 1.	
KW Transcription regulation; Activator; DNA-binding; Nuclear protein.	
FT NON_TER 1 1	
DOMAIN <1 >38 DNA-BINDING (POTENTIAL).	
FT NON_TER 38 38	
SQ SEQUENCE 38 AA; 4313 MW; E790345E375CC9E4 CRC64;	

Query Match 44.0%; Score 40; DB 1; Length 38;  
Best Local Similarity 63.6%; Pred. No. 1.6;  
Matches 7; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

```
QY 1 RRRVVDALNVL 11
   :||:|
Db 26 KRRIYDITNVL 36

RESULT 2
YORA_TTV1
ID_YORA_TTV1 STANDARD; PRT; 70 AA.
AC P19285;
DT 01-NOV-1990 (Rel. 16, Created)
DT 01-NOV-1990 (Rel. 16, Last sequence update)
DE Hypothetical 8.1 kDa protein.
OS Thermoproteus tenax virus 1 (strain KRAL) (TTV1).
OC Viruses; dsDNA viruses, no RNA stage; Lipothirixviridae;
OC Lipothirixvirus.
OX NCBI_TaxID=10480;
RN [1]
RP SEQUENCE FROM N.A.
RA Neumann H.;
RL Submitted (MAR-1989) to the EMBL/GenBank/DBJ databases.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; X14855; CAA32979.1; -.
KW Hypothetical protein.
SQ SEQUENCE 70 AA; 8125 MW; F88F0A8E46323EDC CRC64;

Query Match 35.2%; Score 32; DB 1; Length 70;
Best Local Similarity 37.5%; Pred. No. 79;
Matches 6; Conservative 4; Mismatches 6; Indels 0; Gaps 0;

QY 4 VYDALNVLMAMNISK 19
   |||:|:|:|
Db 32 VYVAITVTLNLARK 47

RESULT 3
FLIE_AQUAE
ID_FLIE_AQUAE STANDARD; PRT; 93 AA.
AC O67242;
DT 30-MAY-2000 (Rel. 39, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Flagellar hook-basal body complex protein flie.
GN FLIE OR AQ_1182.1 OR AQ_1182A.
OS Aquifex aeolicus.
OC Bacteria; Aquificae;
OX NCBI_TaxID=63363;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=VF5;
RX MEDLINE=98196666; PubMed=9537320;
RA Deckert G., Warren P.V., Gaasterland T., Young W.G., Lenox A.L.,
RA Graham D.E., Overbeek R., Snead M.A., Keller M., AuJay M., Huber R.,
RA Feldman R.A., Short J.M., Olson G.J., Swanson R.V.;
RT "The complete genome of the hyperthermophilic bacterium Aquifex
RT aeolicus";
RL Nature 392:353-358(1998).
CC -!- SIMILARITY: BELONGS TO THE FLIE FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; AE000727; AAC07208.1; -.
DR PIR; G70401; G70401.
DR InterPro; IPR001624; Flie.
DR Pfam; PF02049; Flie; 1.
DR PRINTS; PR01006; FLGHOOKFLIE.
DR TIGRFAMs; TIGR00205; flie; 1.
KW Flagella; Complete proteome.
SQ SEQUENCE 93 AA; 10764 MW; 6180AEALC8FB9C39 CRC64;

Query Match 35.2%; Score 32; DB 1; Length 93;
Best Local Similarity 37.5%; Pred. No. 1.1e+02;
Matches 6; Conservative 4; Mismatches 6; Indels 0; Gaps 0;

QY 1 RRRVVDALNVLMAMNI 16
   |:::|:|:|
Db 78 RNKLLEAYNELMRMQV 93

RESULT 4
RL32_ECOLI
ID_RL32_ECOLI STANDARD; PRT; 56 AA.
AC P02435;
DT 21-JUL-1986 (Rel. 01, Created)
DT 21-JUL-1986 (Rel. 01, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE SOS ribosomal protein L32.
DE RPMF OR B1089 OR Z1728 OR ECS1467 OR STM1191 OR STY1230 OR TI729 OR
GN SFI093.
OS Escherichia coli,
OS Escherichia coli O157:H7,
OS Salmonella typhimurium,
OS Salmonella typhi, and
OS Shigella flexneri.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Escherichia.
OX NCBI_TaxID=562, 83334, 602, 601, 623;
RN [1]
RP SEQUENCE FROM N.A.
RC SPECIES=E.coli;
RX MEDLINE=90008815; PubMed=2477362;
RA Tanaka Y., Tsujimura A., Fujita N., Isono S., Isono K.;
RT "Cloning and analysis of an Escherichia coli operon containing the
RT rpmF gene for ribosomal protein L32 and the gene for a 30-kilodalton
RT protein";
RL J. Bacteriol. 171:5707-5712(1989).
RN [2]
RP SEQUENCE FROM N.A.
RC SPECIES=E.coli; STRAIN=KL2 / MG1655;
RX MEDLINE=97426617; PubMed=9278503;
RA Blattner F.R., Plunkett G. III, Bloch C.A., Perna N.T., Burland V.,
RA Riley M., Collado-Vides J., Glasner J.D., Rode C.K., Mayhew G.F.,
RA Gregor J., Davis N.W., Kirkpatrick H.A., Goeden M.A., Rose D.J.,
RA Mau B., Shao Y.;
RT "The complete genome sequence of Escherichia coli K-12.";
RL Science 277:1453-1474(1997).
RN [3]
RP SEQUENCE FROM N.A.
RC SPECIES=E.coli; STRAIN=KL2;
RX MEDLINE=97061202; PubMed=8905232;
RA Oshima T., Aiba H., Baba T., Fujita K., Hayashi K., Honjo A.,
RA Ikemoto K., Inada T., Itoh T., Kajihara M., Kanai K., Kashimoto K.,
RA Kimura S., Kitagawa M., Makino K., Masuda S., Mizobuchi K.,
RA Mori H., Motomura K., Nakamura Y., Nishimoto H., Nishio Y., Saito N.,
RA Sampei G., Seki Y., Tagami H., Takemoto K., Wada C., Yamamoto Y.,
RA Yano M., Horiuchi T.;
RT "A 718-kb DNA sequence of the Escherichia coli K-12 genome
RT corresponding to the 12.7-28.0 min region on the linkage map.";
RL DNA Res. 3:137-155(1996).
RN [4]
RP SEQUENCE FROM N.A.
RC SPECIES=E.coli; STRAIN=O157:H7 / EDL933 / ATCC 700927;
```

RX MEDLINE=21074935; PubMed=11206551;  
 RA Perna N.T., Plunkett G. III, Burland V., Mau B., Glasner J.D.,  
 RA Rose D.J., Mayhew G.F., Evans P.S., Gregor J., Kirkpatrick H.A.,  
 RA Posfai G., Hackett J., Klink S., Boutin A., Shao Y., Miller L.,  
 RA Grobeck E.J., Davis N.W., Lim A., Dimalanta E.T., Potamouis K.,  
 RA Apodaca J., Anantharaman T.S., Lin J., Yen G., Schwartz D.C.,  
 RA Welch R.A., Blattner F.R.;  
 RT "Genome sequence of enterohaemorrhagic *Escherichia coli* O157:H7";  
 RL Nature 409:529-533(2001).  
 RN [5]  
 RN SEQUENCE FROM N.A.  
 RC SPECIES=*E. coli*; STRAIN=O157:H7 / RIMD 0509952;  
 RX MEDLINE=21156231; PubMed=11258796;  
 RA Hayashi T., Makino K., Onishi M., Kurokawa K., Ishii K., Yokoyama K.,  
 RA Han C.-G., Ohtsubo E., Nakayama K., Murata T., Tanaka M., Tobe T.,  
 RA Iida T., Takami H., Honda T., Sasakawa C., Ogasawara N., Yasunaga T.,  
 RA Kuhara S., Shiba T., Hattori M., Shinagawa H.;  
 RT "Complete genome sequence of enterohemorrhagic *Escherichia coli*  
 RT O157:H7 and genomic comparison with a laboratory strain K-12";  
 RL DNA Res. 8:11-22(2001).  
 RN [6]  
 RN SEQUENCE.  
 RC SPECIES=*E. coli*; STRAIN=K;  
 RX MEDLINE=76119562; PubMed=765258;  
 RA Wittmann-Liebold B., Greuer B., Pannenbecker R.;  
 RT "The primary structure of protein L32 from the 50S subunit of  
 RT *Escherichia coli* ribosomes";  
 RL Hoppe-Seyler's Z. Physiol. Chem. 356:1977-1979(1975).  
 RN [7]  
 RN SEQUENCE.  
 RC SPECIES=*E. coli*; STRAIN=MRE-600;  
 RA Vinokurov L.M., Alakhov Y.B., Golov E.A., Ovchinnikov Y.A.;  
 RT "The primary structure of ribosomal protein L32 from *E. coli* MRE-600  
 RT ribosomes";  
 RL Bioorg. Khim. 2:1013-1017(1976).  
 RN [8]  
 RN SEQUENCE FROM N.A.  
 RC SPECIES=*S. typhimurium*; STRAIN=LT2;  
 RX MEDLINE=98317265; PubMed=9642179;  
 RA Zhang Y., Cronan J.E. Jr.;  
 RT "Transcriptional analysis of essential genes of the *Escherichia coli*  
 RT fatty acid biosynthesis gene cluster by functional replacement with  
 RT the analogous *Salmonella typhimurium* gene cluster";  
 RL J. Bacteriol. 180:3295-3303(1998).  
 RN [9]  
 RN SEQUENCE FROM N.A.  
 RC SPECIES=*S. typhimurium*; STRAIN=LT2 / SGSC1412 / ATCC 700720;  
 RX MEDLINE=21534948; PubMed=11677609;  
 RA McClelland M., Sanderson K.E., Spieth J., Clifton S.W., Latreille P.,  
 RA Courtney L., Porwollik S., Ali J., Dante M., Du F., Hou S., Layman D.,  
 RA Leonard S., Nguyen C., Scott K., Holmes A., Grewal N., Mulvaney E.,  
 RA Ryan E., Sun H., Flores L., Miller W., Stoneking T., Nhan M.,  
 RA Waterston R., Wilson R.K.;  
 RT "Complete genome sequence of *Salmonella enterica* serovar *Typhimurium*  
 RT LT2";  
 RL Nature 413:852-856(2001).  
 RN [10]  
 RN SEQUENCE FROM N.A.  
 RC SPECIES=*S. typhi*; STRAIN=CT18;  
 RX MEDLINE=21534947; PubMed=11677608;  
 RA Parkhill J., Dougan G., James K.D., Thomson N.R., Pickard D., Wain J.,  
 RA Churcher C., Mungall K.L., Bentley S.D., Holden M.T.G., Sebaihia M.,  
 RA Baker S., Basham D., Brooks K., Chillingworth T., Connor P.,  
 RA Cronin A., Davis P., Davies R.M., Dowd L., White N., Farrar J.,  
 RA Feltwell T., Hamlin N., Haque A., Hien T.T., Holroyd S., Jagels K.,  
 RA Krogh A., Larsen T.S., Leather S., Moule S., O'Garra P., Parry C.,  
 RA Quail M., Rutherford K., Simmonds M., Skelton J., Stevens K.,  
 RA Whitehead S., Barrett B.G.;  
 RT "Complete genome sequence of a multiple drug resistant *Salmonella*  
 RT *enterica* serovar *Typhi* CT18";  
 RL Nature 413:848-852(2001).  
 RN [11]  
 RN SEQUENCE FROM N.A.

RC SPECIES=*S. typhi*; STRAIN=Ty2 / ATCC 700931;  
 RX MEDLINE=22531367; PubMed=12644504;  
 RA Deng W., Liou S.-R., Plunkett G. III, Mayhew G.F., Rose D.J.,  
 RA Burland V., Kodyanani V., Schwartz D.C., Blattner F.R.;  
 RT "Comparative genomics of *Salmonella enterica* serovar *Typhi* strains Ty2  
 RT and CT18";  
 RL J. Bacteriol. 185:2330-2337(2003).  
 RN [12]  
 RN SEQUENCE FROM N.A.  
 RC SPECIES=*S. flexneri*; STRAIN=301 / Serotype 2a;  
 RX MEDLINE=22272406; PubMed=12384590;  
 RA Jin Q., Yuan Z., Xu J., Wang Y., Shen Y., Lu W., Wang J., Liu H.,  
 RA Yang J., Yang F., Zhang X., Zhang J., Yang G., Wu H., Qu D., Dong J.,  
 RA Sun L., Xue Y., Zhao A., Gao Y., Zhu J., Kan B., Ding K., Chen S.,  
 RA Cheng H., Yao Z., He B., Chen R., Ma D., Qiang B., Wen Y., Hou Y.,  
 RA Yu J.;  
 RT "Genome sequence of *Shigella flexneri* 2a: insights into pathogenicity  
 RT through comparison with genomes of *Escherichia coli* K12 and O157";  
 RL Nucleic Acids Res. 30:4432-4441(2002).  
 RN [13]  
 RN MASS SPECTROMETRY.  
 RC SPECIES=*E. coli*; STRAIN=K12 / ATCC 25404;  
 RX MEDLINE=99196679; PubMed=10094780;  
 RA Arnold R.J., Reilly J.P.;  
 RT "Observation of *Escherichia coli* ribosomal proteins and their  
 RT posttranslational modifications by mass spectrometry";  
 RL Anal. Biochem. 269:105-112(1999).  
 CC -1- MASS SPECTROMETRY; MW=6315.1; METHOD=MALDI.  
 CC -1- SIMILARITY: BELONGS TO THE L32P FAMILY OF RIBOSOMAL PROTEINS.  
 CC -----  
 CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 CC ENBL; M29698; AAA24575.1; -;  
 DR ENBL; AE000209; AAC74173.1; -;  
 DR ENBL; D90744; BAA35897.1; -;  
 DR ENBL; A8005319; AAG55835.1; -;  
 DR ENBL; AP002555; BAA34890.1; -;  
 DR ENBL; AF044668; AAC38646.1; -;  
 DR ENBL; AE008751; AAL20120.1; -;  
 DR ENBL; AL627269; CAD08315.1; -;  
 DR ENBL; AE016839; AAO69353.1; -;  
 DR ENBL; AB015136; AAN42712.1; -;  
 DR PIR; C90812; C90812.  
 DR PIR; G85671; G85671.  
 DR PIR; JY0048; R5EC32.  
 DR ECO2DBASE; I011.7; 6TH EDITION.  
 DR EcoGene; EG10890; rpmf.  
 DR StyGene; SGI0698; rpmf.  
 DR HAMAP; MF\_00340; -; 1.  
 DR InterPro; IPR005718; S32\_bact.org.  
 DR TIGRFAMs; TIGR01031; rpmf\_bact; 1.  
 KW Ribosomal protein; Complete proteome.  
 FT INIT\_MET 0 0  
 SQ SSSEQUENCE 56 AA; 6315 MW; B7A9510E78E4E94 CRC64;  
  
 Query Match 34.1%; Score 31; DB 1; Length 56;  
 Best Local Similarity 33.3%; Pred. No. 93;  
 Matches 5; Conservative 6; Mismatches 4; Indels 0; Gaps 0;  
  
 Qy 2 RRYVDALNVLMMNVI 16  
 ||| :||| : : : :  
 Db 15 RRSHDALTAVTSLSV 29  
  
 RESULT 5  
 PLP\_CRYNE STANDARD; PRT; 79 AA.  
 ID - PLP\_CRYNE

AC	P82476;	DR	EMBL; U75930; AAC59075.1; -.
DT	30-MAY-2000 (Rel. 39, Created)	KW	Hypothetical protein.
DT	30-MAY-2000 (Rel. 39, Last sequence update)	SQ	SEQUENCE 84 AA; 9427 MW; 46F6656120231346 CRC64;
DT	16-OCT-2001 (Rel. 40, Last annotation update)		
DE	Extracellular phospholipase (EC 3.1.1.5) (Fragments).	Query Match	34.1%; Score 31; DB 1; Length 84;
OS	Cryptococcus neoformans (Filobasidiella neoformans).	Best Local Similarity	31.6%; Pred. No. 1.4e+02;
OC	Eukaryota; Fungi; Basidiomycota; Hymenomycetes; Heterobasidiomycetes;	Matches	6; Conservative 7; Mismatches 4; Indels 2; Gaps 1;
OC	Tremellomycetidae; Tremellales; Tremellaceae; Filobasidiella.		
OX	NCBI_TaxID=5207;	QY	1 RRRVYDA--LNVLMAMNII 17
RN	[1]	DB	56 RRSFYESLKLNCIVCINVL 74
RP	SEQUENCE		
RC	STRAIN=Isolate BL-1;		
RC	MEDLINE=21060447; PubMed=10749672;		
RA	Chen S.C.A., Wright L.C., Golding J.C., Sorrell T.C.;		
RT	"Purification and characterization of secretory phospholipase B,		
RT	lysophospholipase and lysophospholipase/transacylase from a virulent		
RT	strain of the pathogenic fungus Cryptococcus neoformans.";		
RL	Biochem. J. 347:431-439 (2000).		
CC	-1- FUNCTION. A NOVEL FUNGAL PROTEIN THAT EXHIBITS PHOSPHOLIPASE B		
CC	(PLB), LYSOPHOSPHOLIPASE (LPL) AND LYSOPHOSPHOLIPASE/TRANSACYLASE		
CC	(LPL) ACTIVITIES. ACTIVE ONLY AT ACIDIC PH.		
CC	-1- CATALYTIC ACTIVITY: 2-lysophosphatidylcholine + H(2)O =		
CC	glycerophosphocholine + a fatty acid anion.		
CC	-1- ENZYME REGULATION: INHIBITED BY FERRIC ION.		
CC	-1- SUBCELLULAR LOCATION: Secreted.		
KW	Hydrolase; Lipid degradation; Glycoprotein.		
FT	NON_CONS 21		
FT	NON_CONS 40		
FT	NON_CONS 56		
FT	NON_CONS 65		
FT	NON_CONS 73		
FT	NON_TER 79		
FT	NON_TER 79		
SQ	SEQUENCE 79 AA; 8637 MW; 685F80B144195AF2 CRC64;		
		Query Match	34.1%; Score 31; DB 1; Length 79;
		Best Local Similarity	46.7%; Pred. No. 1.3e+02;
		Matches	7; Conservative 2; Mismatches 6; Indels 0; Gaps 0;
QY	1 RRRVYDALNVLMAMN 15		
DB	39 RRVADAFNMEXYLN 53		
		RESULT 6	
Y073_NPVOP	STANDARD; PRT; 84 AA.		
ID	O10326;		
DT	01-NOV-1997 (Rel. 35, Created)		
DT	01-NOV-1997 (Rel. 35, Last sequence update)		
DT	01-NOV-1997 (Rel. 35, Last annotation update)		
DE	Hypothetical 9.4 kDa protein (ORF76)		
OS	Oryza pseudotsugata multicapsid polyhedrosis virus (OpNPV).		
OC	Viruses; dsDNA viruses, no RNA stage; Baculoviridae;		
OC	Nucleopolyhedrovirus.		
OX	NCBI_TaxID=164623;		
RN	[1]		
RP	SEQUENCE FROM N.A.		
RP	MEDLINE=97271300; PubMed=9126251;		
RA	Ahrens C.H., Russell R.R., Funk C.J., Evans J., Harwood S.,		
RA	Rohrmann G.F.;		
RT	"The sequence of the Orzyia pseudotsugata multinucleocapsid nuclear		
RT	polyhedrosis virus genome.";		
RL	Virology 229:381-399 (1997).		
CC	-1- SIMILARITY: TO CORRESPONDING ORF IN ACNPV.		
CC			
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration		
CC	between the Swiss Institute of Bioinformatics and the EMBL outstation -		
CC	the European Bioinformatics Institute. There are no restrictions on its		
CC	use by non-profit institutions as long as its content is in no way		
CC	modified and this statement is not removed. Usage by and for commercial		
CC	entities requires a license agreement (See <a href="http://www.isb-sib.ch/announce/">http://www.isb-sib.ch/announce/</a>		
CC	or send an email to <a href="mailto:license@isb-sib.ch">license@isb-sib.ch</a> ).		
CC			
		Query Match	33.5%; Score 30.5; DB 1; Length 77;
		Best Local Similarity	39.1%; Pred. No. 1.6e+02;
		Matches	9; Conservative 2; Mismatches 3; Indels 9; Gaps 1;
QY	6 DALNV-----LMAMNIISK 19		
DB	23 DNLNIELKSLGIDLSAMNLIK 45		
		RESULT 8	
ATP8_CANPA	STANDARD; PRT; 48 AA.		
ID	-ATP8_CANPA		
AC	P17345;		



```

DT 01-AUG-1990 (Rel. 15, Created)
DT 01-AUG-1990 (Rel. 15, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE ATP synthase protein 8 (EC 3.6.3.14) (ATPase subunit 8) (A6L).
GN ATP8 OR AAPI.
OS Candida parapsilosis (Yeast).
OC Mitochondrion.
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; mitosporic Saccharomycetales; Candida.
OX NCBI_TaxID=5480;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CBS 7154 / SP1;
RX MEDLINE=90332440; PubMed=2143015;
RA Guelin E., Velours J., Guerin M.;
RT "Cloning and sequencing of a fragment of the linear mitochondrial DNA
RT of the yeast Candida parapsilosis supporting genes encoding subunit 8
RT of F0 ATP synthase and a putative t-RNA(Pro).";
RL Nucleic Acids Res. 18:4267-4267(1990).
CC -1- FUNCTION: THIS IS ONE OF THE CHAINS OF THE NONENZYMATIC COMPONENT
CC (CF(0) SUBUNIT) OF THE MITOCHONDRIAL ATPASE COMPLEX.
CC -1- CATALYTIC ACTIVITY: ATP + H(2)O + H(+) (In) = ADP + phosphate +
CC H(+) (Out).
CC -1- SUBCELLULAR LOCATION: Membrane-bound.
CC -1- SIMILARITY: BELONGS TO THE ATPASE PROTEIN 8 FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: X52115; CAA36361.1; -.
DR PIR: S10465; PWCK8P.
KW Hydrogen ion transport; CF(0); Mitochondrion; Transmembrane.
FT TRANSMEM 12 32 POTENTIAL.
SQ SEQUENCE 48 AA; 5492 MW; 0E7876341CAF0747 CRC64;
-----
Query Match 33.0%; Score 30; DB 1; Length 48;
Best Local Similarity 58.3%; Pred. No. 1.2e+02;
Matches 7; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 8 LNVLMNMNIISK 19
DB 36 LRLLIARNIIK 47
| : | : | : | : |
| : | : | : | : |

RESULT 9
SCP2 MESMA
ID _SCP2 MESMA STANDARD; PRT; 56 AA.
AC QSNJ7; P58491;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE Neurotoxin Bmp03 precursor (Potassium ion channel blocker P02P).
OS Mesobuthus martensii (Manchurian scorpion) (Buthus martensii).
OC Eukaryota; Metazoa; Arthropoda; Chelicerata; Arachnida; Scorpiones;
OC Buthoidea; Buthidae; Mesobuthus.
OX NCBI_TaxID=34649;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Venom gland;
RX MEDLINE=99402983; PubMed=10471839;
RA Zhu S.-Y., Li W.-X., Zeng X.-C., Jiang D.-H., Mao X., Liu H.;
RT "Molecular cloning and sequencing of two 'short chain' and two 'long
RT chain' K(+) channel-blocking peptides from the Chinese scorpion Buthus
RT martensii Karsch.";
RL FEBS Lett. 457:509-514(1999).
RN [2]
RP STRUCTURE BY NMR OF 29-56.
RC TISSUE=Venom;

```

```

RX MEDLINE=20530297; PubMed=11076505;
RA Xu Y.Q., Wu J.H., Pei J.M., Shi Y.Y., Ji Y.H., Tong Q.C.;
RT "Solution structure of Bmp02, a new potassium channel blocker from the
RT venom of the Chinese scorpion Buthus martensii Karsch.";
RL Biochemistry 39:13669-13675(2000).
CC -1- SUBCELLULAR LOCATION: Secreted.
CC -1- TISSUE SPECIFICITY: Expressed by the venom gland.
CC -1- SIMILARITY: BELONGS TO THE SCORPION LEIUROTOXIN FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: AF132975; AAF31296.1; -.
DR PDB: 1D09; 28-MAR-01.
KW Toxin; Neurotoxin; Ionic channel inhibitor;
KW Potassium channel inhibitor; Signal; 3D-structure.
FT SIGNAL 1 28
FT CHAIN 29 56 NEUROTOXIN BMP02.
FT DISULFID 31 47
FT DISULFID 34 52
FT DISULFID 38 54
SQ SEQUENCE 56 AA; 6015 MW; 70953032042F8672 CRC64;
-----
Query Match 33.0%; Score 30; DB 1; Length 56;
Best Local Similarity 26.7%; Pred. No. 1.4e+02;
Matches 4; Conservative 8; Mismatches 3; Indels 0; Gaps 0;

QY 3 RVYDALNVLMNMNI 17
DB 3 RLFTLVILVLMNMV 17
| : | : | : | : |
| : | : | : | : |

RESULT 10
SCP3 MESMA
ID _SCP3 MESMA STANDARD; PRT; 56 AA.
AC Q08DL; 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Neurotoxin Bmp03 precursor (Potassium ion channel blocker P03).
OS Mesobuthus martensii (Manchurian scorpion) (Buthus martensii).
OC Eukaryota; Metazoa; Arthropoda; Chelicerata; Arachnida; Scorpiones;
OC Buthoidea; Buthidae; Mesobuthus.
OX NCBI_TaxID=34649;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Venom gland;
RX MEDLINE=99313193; PubMed=10386622;
RA Wu J.-J., Dai L., Lan Z.-D., Chi C.-W.;
RT "Genomic organization of three neurotoxins active on small conductance
RT Ca2+-activated potassium channels from the scorpion Buthus martensii
RT Karsch.";
RL FEBS Lett. 452:360-364(1999).
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=Venom gland;
RX MEDLINE=99402983; PubMed=10471839;
RA Zhu S.-Y., Li W.-X., Zeng X.-C., Jiang D.-H., Mao X., Liu H.;
RT "Molecular cloning and sequencing of two 'short chain' and two 'long
RT chain' K(+) channel-blocking peptides from the Chinese scorpion Buthus
RT martensii Karsch.";
RL FEBS Lett. 457:509-514(1999).
RN [2]
RP STRUCTURE BY NMR OF 29-56.
RC TISSUE=Venom;

```

CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC  
DR EMBL; AF097408; AF01253.1; -  
DR EMBL; AF156170; AF29463.1; -  
DR HSSP; Q9NJ7; IDU9.  
KW Toxin; Neurotoxin; Ionic channel inhibitor;  
KW Potassium channel inhibitor; Signal.  
FT SIGNAL 1 28 BY SIMILARITY.  
FT CHAIN 29 56 NEUROTOXIN EMP03.  
FT DISULFID 31 47 BY SIMILARITY.  
FT DISULFID 34 52 BY SIMILARITY.  
FT DISULFID 38 54 BY SIMILARITY.  
SQ SEQUENCE 56 AA; 6001 MW; 70953032007E8672 CRC64;  
Query Match 33.0%; Score 30; DB 1; Length 56;  
Best Local Similarity 26.7%; Pred. No. 1.4e+02;  
Matches 4; Conservative 8; Mismatches 3; Indels 0; Gaps 0;  
QY 3 RYVDALNVLMMNII 17  
Db 3 RLFTLVILVLMNMV 17  
RESULT 11  
VREP BPPHH STANDARD; PRT; 89 AA.  
AC P22562;  
DT 01-AUG-1991 (Rel. 19, Created)  
DT 01-AUG-1991 (Rel. 19, Last sequence update)  
DT 01-AUG-1991 (Rel. 19, Last annotation update)  
DE Putative repressor.  
GN T6.  
OS Bacteriophage phi-H.  
OC Viruses; dsDNA viruses, no RNA stage; Caudovirales; Podoviridae.  
OX NCBI\_TaxID=10771;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=91123222; PubMed=1991733;  
RA Ken R., Hackett N.R.;  
RT "Halobacterium halobium strains lysogenic for phage phi H contain a  
RT protein resembling coliphage repressors.";  
RL J. Bacteriol. 173:955-960(1991).  
CC -1- FUNCTION: POSSIBLY PREVENTS RNA POLYMERASE ACCESS TO THE PROMOTERS  
CC FOR LYTIC CELL CYCLE TRANSCRIPTION.  
CC -1- SIMILARITY: TO THE HELIX-TURN-HELIX FAMILY OF DNA-BINDING  
CC PROTEINS, AND IN PARTICULAR TO COLIPHAGE REPRESSORS.  
CC  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC  
DR EMBL; X52504; CAA36746.1; -  
DR PIR; A38167; A38167  
KW Repressor; DNA-binding; Transcription regulation.  
FT DNA BIND 29 38 H-T-H MOTIF.  
SQ SEQUENCE 89 AA; 10167 MW; 3D21EDB53F01741 CRC64;  
Query Match 33.0%; Score 30; DB 1; Length 89;  
Best Local Similarity 35.3%; Pred. No. 2.3e+02;  
Matches 6; Conservative 4; Mismatches 7; Indels 0; Gaps 0;  
QY 1 RRRVDALNVLMMNII 17

Db 40 RRRISDRCTVLVDRLGL 56  
RESULT 12  
YFRC PROVV STANDARD; PRT; 96 AA.  
AC P20927;  
DT 01-FEB-1991 (Rel. 17, Created)  
DT 01-FEB-1991 (Rel. 17, Last sequence update)  
DT 01-FEB-1995 (Rel. 31, Last annotation update)  
DE PRD operon hypothetical protein C.  
OS Proteus vulgaris.  
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;  
OC Enterobacteriaceae; Proteus.  
OX NCBI\_TaxID=585;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=88004470; PubMed=3308458;  
RA Cole S.T.;  
RT "Nucleotide sequence and comparative analysis of the frd operon  
RT encoding the fumarate reductase of Proteus vulgaris. Extensive  
RT sequence divergence of the membrane anchors and absence of an  
RT frd-linked ampC cephalosporinase gene.";  
RL Eur. J. Biochem. 167:481-488(1987).  
CC -1- SIMILARITY: BELONGS TO THE HUPF/HYPC FAMILY.  
CC  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC  
CC EMBL; X06151; CAA29511.1; -  
DR PIR; S00119; S00119.  
DR InterPro; IPR001109; HupF\_HypC.  
DR Pfam; PF01455; HupF\_HypC; 1.  
DR PIRSF; PIRSF005618; HupF\_HypC; 1.  
DR PRINTS; PR00445; HUPFHYP.  
DR PRODOM; PD003112; HupF\_HypC; 1.  
DR TIGRFAMs; TIGR00074; hupF\_hupF; 1.  
DR PROSITE; PS01097; HUPF\_HYP; 1.  
KW Hypothetical protein.  
SQ SEQUENCE 96 AA; 10300 MW; EC094F1F37956EE3 CRC64;  
Query Match 33.0%; Score 30; DB 1; Length 96;  
Best Local Similarity 66.7%; Pred. No. 2.5e+02;  
Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;  
QY 6 DALNVLMM 14  
Db 67 ETNLALMM 75  
RESULT 13  
APR HUMAN STANDARD; PRT; 54 AA.  
AC Q13794;  
DT 30-MAY-2000 (Rel. 39, Created)  
DT 30-MAY-2000 (Rel. 39, Last sequence update)  
DT 28-FEB-2003 (Rel. 41, Last annotation update)  
DE Phorbol-12-myristate-13-acetate-induced protein 1 (PMA-induced  
DE protein 1) (Immediate-early-response protein APR).  
GN PMAIP1.  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
OX NCBI\_TaxID=9606;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=90376412; PubMed=2398525;





```

KW Oxidoreductase; NAD; Ubiquinone; Mitochondrion.
SQ SEQUENCE 98 AA; 10840 MW; D0FF9BC309048774 CRC64;

Query Match 31.8%; Score 29; DB 1; Length 98;
Best Local Similarity 55.6%; Pred. No. 3.8e+02;
Matches 5; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 8 LNLVLMWNI 16
   ||::||:|
Db 6 LNIIMAFSI 14

RESULT 19
COXG THUOB STANDARD; PRT; 34 AA.
AC P80976;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Cytochrome c oxidase polypeptide VIB (EC 1.9.3.1) (Fragments).
OS Thunnus obesus (Bigeye tuna).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
OC Acanthomorpha; Acanthopterygii; Perciformes; Scombroidei;
OC Scombridae; Thunnus.
OX NCBI_TaxID=8241;
RN [1]
RP SEQUENCE.
RC TISSUE=Heart, and Liver;
RX MEDLINE=97454291; PubMed=9310366;
RA Arnold S., Lee I., Kim M., Song E., Linder D., Lottspeich F.,
RA Kadenbach B.;
RT "The subunit structure of cytochrome-c oxidase from tuna heart and
RT liver.";
RL Eur. J. Biochem. 248:99-103 (1997).
CC -1- FUNCTION: THIS PROTEIN IS ONE OF THE NUCLEAR-CODED POLYPEPTIDE
CC CHAINS OF CYTOCHROME C OXIDASE, THE TERMINAL OXIDASE IN
CC MITOCHONDRIAL ELECTRON TRANSPORT. THIS PROTEIN MAY BE ONE OF THE
CC HEME-BINDING SUBUNITS OF THE OXIDASE.
CC -1- CATALYTIC ACTIVITY: 4 ferrocyclochrome c + O(2) = 4 ferricytochrome
CC c + 2 H(2)O.
CC -1- SIMILARITY: BELONGS TO THE CYTOCHROME C OXIDASE VIB FAMILY.
CC -1- CAUTION: THE ORDER OF THE PEPTIDES SHOWN IS UNKNOWN.
KW Oxidoreductase; Mitochondrion.
FT NON_TER 1 1
FT NON_CONS 15 16
FT NON_TER 34 34
FT NON_TER 34 34
SQ SEQUENCE 34 AA; 4035 MW; B34A390BA1F05546 CRC64;

Query Match 30.8%; Score 28; DB 1; Length 34;
Best Local Similarity 85.7%; Pred. No. 1.8e+02;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 RRVYDAL 8
   |||||
Db 12 RRVYKAL 18

RESULT 20
RETS BOVIN STANDARD; PRT; 42 AA.
AC P82708;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Retinol-binding protein III, cellular (CRBP-III) (Fragment).
GN RBP5.
OS Bos taurus (Bovine).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
OC Bovidae; Bovinae; Bos.
OX NCBI_TaxID=9913;
RN [1]

```

```

RP SEQUENCE, AND FUNCTION.
RC TISSUE=Kidney;
RX MEDLINE=21173623; PubMed=11274389;
RA Polli C., Calderone V., Ottonello S., Bolchi A., Zanotti G.,
RA Stoppini M., Berni R.;
RT "Identification, retinoid binding and X-ray analysis of a human
RT retinol-binding protein.";
RL Proc. Natl. Acad. Sci. U.S.A. 98:3710-3715 (2001).
CC -1- FUNCTION: Intracellular transport of retinol.
CC -1- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
CC -1- TISSUE SPECIFICITY: Kidney.
CC -1- SIMILARITY: BELONGS TO THE FABP/P2/CRBP/CRABP FAMILY OF
CC TRANSPORTERS.
DR HSP; P82980; IGGL.
DR InterPro; IPR000463; Fatty acid BP.
DR InterPro; IPR000566; Lipocalin_cytFABP.
DR Pfam; PF00061; Lipocalin_1.
DR PROSITE; PS00214; FABP; FALSE_NEG.
KW Vitamin A; Retinol-binding; Transport.
FT NON_TER 42 42
SQ SEQUENCE 42 AA; 4892 MW; ACB4F1399PDD7F09 CRC64;

Query Match 30.8%; Score 28; DB 1; Length 42;
Best Local Similarity 58.3%; Pred. No. 2.3e+02;
Matches 7; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 7 ALNVLMMNIIS 18
   |||||
Db 22 ALNVLMMALRKIA 33

RESULT 21
RL32 MYCPU STANDARD; PRT; 61 AA.
AC Q98QW7;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE 50S ribosomal protein L32.
GN RPLP OR MYPV 3240.
OS Mycoplasma pulmonis.
OC Bacteria; Firmicutes; Mollicutes; Mycoplasmataceae; Mycoplasma.
OX NCBI_TaxID=2107;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=UAB CTIP;
RX MEDLINE=21267165; PubMed=11353084;
RA Chambaud I., Heilig R., Ferris S., Barbe V., Samson D., Galisson F.,
RA Moszer I., Dybvig K., Wroblewski H., Viari A., Rocha E.P.C.,
RA Blanchard A.;
RT "The complete genome sequence of the murine respiratory pathogen
RT Mycoplasma pulmonis.";
RL Nucleic Acids Res. 29:2145-2153 (2001).
CC -1- SIMILARITY: BELONGS TO THE L32P FAMILY OF RIBOSOMAL PROTEINS.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; AL445564; CAC13497.1; -.
CC FBL; D90552; D90552.
CC MypuList; MYPV_3240; -.
CC HAMAP; MF 00340; -.
CC InterPro; IPR002677; Ribosomal_L32p.
CC InterPro; IPR005718; S32 bact Org.
CC Pfam; PF01783; Ribosomal_L32p; 1.
CC TIGRFAMs; TIGR01031; rpmF_bact; 1.
CC Ribosomal protein; Complete proteome.
SQ SEQUENCE 61 AA; 6904 MW; B7FD0E475ACA3DB1 CRC64;

```

RP	SEQUENCE FROM N.A.				
RC	SPECIES=S flexneri; STRAIN=301 / Serotype 2a;				
RX	MEDLINE=22272406; PubMed=12384590;				
RA	Jin Q., Yuan Z., Xu J., Wang Y., Shen Y., Lu W., Wang J., Liu H.,				
RA	Yang J., Yang F., Zhang X., Zhang J., Yang G., Wu H., Qu D., Dong J.,				
RA	Sun L., Xue Y., Zhao A., Gao Y., Zhu J., Kan B., Ding K., Chen S.,				
RA	Cheng H., Yao Z., He B., Chen R., Ma D., Qiang B., Wen Y., Hou Y.,				
RA	Yu J.;				
RT	"Genome sequence of Shigella flexneri 2a: insights into pathogenicity				
RT	through comparison with genomes of Escherichia coli K12 and O157.";				
RL	Nucleic Acids Res. 30:4432-4441 (2002).				
[7]					
RP	IDENTIFICATION.				
RC	SPECIES=E.coli;				
RX	MEDLINE=95075659; PubMed=7984428;				
RA	Borodovsky M., Rudd K.E., Koonin E.V.;				
RT	"Intrinsic and extrinsic approaches for detecting genes in a				
RT	bacterial genome.";				
RL	Nucleic Acids Res. 22:4756-4767 (1994).				
CC	- - SUBCELLULAR LOCATION: Attached to the membrane by a lipid anchor				
CC	(Potential).				
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration				
CC	between the Swiss Institute of Bioinformatics and the EMBL outstation -				
CC	the European Bioinformatics Institute. There are no restrictions on its				
CC	use by non-profit institutions as long as its content is in no way				
CC	modified and this statement is not removed. Usage by and for commercial				
CC	entities requires a license agreement (See http://www.isb-sib.ch/announce/				
CC	or send an email to license@isb-sib.ch).				
CC	-----				
EMBL; M87049; -; NOT ANNOTATED_CDS.					
DR	EMBL; AE000457; -; NOT ANNOTATED_CDS.				
DR	EMBL; X66782; -; NOT ANNOTATED_CDS.				
DR	EMBL; AE016769; AAN83162.1; -;				
DR	EMBL; AE005612; AAGS9001.1; -;				
DR	EMBL; AP002567; -; NOT ANNOTATED_CDS.				
DR	EMBL; AE015395; AAN45322.1; -;				
DR	PIR; E86067; E86067.				
DR	Ecogene; EG12353; YifL.				
DR	PROSITE; PS00013; PROKAR LIPOPROTEIN; 1.				
KW	Hypothetical protein; Membrane; Lipoprotein; Signal;				
KW	Complete proteome.				
FT	SIGNAL 1 19 POTENTIAL.				
FT	CHAIN 20 67 HYPOTHETICAL LIPOPROTEIN YIFL.				
FT	LIPID 20 20 N-ACYL DIGLYCERIDE (POTENTIAL).				
FT	CONFLICT 60 67 DGPSQVNY -> MVHPR (IN REF. 2).				
SQ	SEQUENCE 67 AA; 7177 MW; 4A52CF48A55A7F9 CRC64;				
	Query Match 30.8%; Score 28; DB 1; Length 67;				
	Best Local Similarity 38.5%; Pred. No. 3.8e+02;				
	Matches 5; Conservative 4; Mismatches 4; Indels 0; Gaps 0;				
QY	2 RRVYDALNVLMAM 14				
Db	2 KNVFKALTVLLTL 14				
RESULT 23					
REPL_ECOLI					
ID	REPL_ECOLI	STANDARD;	PRT;	77	AA.
AC	P10396;				
DT	01-MAR-1989 (Rel. 10, Created)				
DT	01-FEB-1996 (Rel. 33, Last sequence update)				
DT	15-DEC-1998 (Rel. 37, Last annotation update)				
DE	Replication initiation protein (Fragment).				
GN	REPA OR REPAL.				
OS	Escherichia coli.				
OG	Plasmid IncF1 ColV2-K94.				
OC	Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;				
OC	Enterobacteriaceae; Escherichia.				
OX	NCBI_TaxID=562;				
RN	[1]				

Query Match	30.8%;	Score 28;	DB 1;	Length 61;	
Best Local Similarity	50.0%;	Pred. No. 3.4e+02;			
Matches	5;	Conservative	3;	Mismatches 2; Indels 0; Gaps 0;	
QY	1 RRVYDALNV 10	: : :			
Db	16 KRQTHDALKV 25				
RESULT 22					
YIFL_ECOLI					
ID	YIFL_ECOLI	STANDARD;	PRT;	67	AA.
AC	P39166; Q8X3Y5;				
DT	01-FEB-1995 (Rel. 31, Created)				
DT	01-FEB-1995 (Rel. 31, Last sequence update)				
DT	28-FEB-2003 (Rel. 41, Last annotation update)				
DE	Hypothetical lipoprotein yifL precursor.				
GN	YIFL OR B3808.1 OR C4729 OR ECS4737.1 OR SF3886.				
OS	Escherichia coli.				
OS	Escherichia coli O6				
OS	Escherichia coli O157:H7, and				
OS	Shigella flexneri.				
OC	Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;				
OC	Enterobacteriaceae; Escherichia.				
OX	NCBI_TaxID=562, 217992, 83334, 623;				
[1]					
RP	SEQUENCE FROM N.A.				
RC	SPECIES=E.coli; STRAIN=K12 / MGI655;				
RX	MEDLINE=92358234; PubMed=1379743;				
RA	Daniels D.L., Plunkett G. III, Burland V.D., Blattner F.R.;				
RT	"Analysis of the Escherichia coli genome: DNA sequence of the region				
RT	from 84.5 to 86.5 minutes."				
RL	Science 257:771-778 (1992).				
[2]					
RP	SEQUENCE FROM N.A.				
RC	SPECIES=E.coli; STRAIN=K12;				
RA	Glaser P., Sismeiro O., Danchin A.;				
RL	Submitted (JUN-1992) to the EMBL/GenBank/DBJ databases.				
[3]					
RP	SEQUENCE FROM N.A.				
RC	SPECIES=E.coli; STRAIN=O6:H1 / CFT073 / ATCC 700928;				
RX	MEDLINE=22388234; PubMed=12471157;				
RA	Welch R.A., Burland V., Plunkett G. III, Redford P., Roesch P.,				
RA	Rasko D., Buckles E.L., Liou S.-R., Boutin A., Hackett J., Stroud D.,				
RA	Mayhew G.F., Rose D.J., Zhou S., Schwartz D.C., Perna N.T.,				
RA	Mobley H.L.T., Donnenberg M.S., Blattner F.R.;				
RT	"Extensive mosaic structure revealed by the complete genome sequence				
RT	of uropathogenic Escherichia coli;"				
RL	Proc. Natl. Acad. Sci. U.S.A. 99:17020-17024 (2002).				
[4]					
RP	SEQUENCE FROM N.A.				
RC	SPECIES=E.coli; STRAIN=O157:H7 / EDL933 / ATCC 700927;				
RX	MEDLINE=21074935; PubMed=11206551;				
RA	Perna N.T., Plunkett G. III, Burland V., Mau B., Glasner J.D.,				
RA	Rose D.J., Mayhew G.F., Evans P.S., Gregor J., Kirkpatrick H.A.,				
RA	Postai G., Hackett J., Klink S., Boutin A., Shao Y., Miller L.,				
RA	Grotbeck E.J., Davis N.W., Lim A., Dimalanta E.T., Potamousis K.,				
RA	Apodaca J., Anantharaman T.S., Lin J., Yen G., Schwartz D.C.,				
RA	Welch R.A., Blattner F.R.;				
RT	"Genome sequence of enterohaemorrhagic Escherichia coli O157:H7;"				
RL	Nature 409:529-533 (2001).				
[5]					
RP	SEQUENCE FROM N.A.				
RC	SPECIES=E.coli; STRAIN=O157:H7 / RIMD 0509952;				
RX	MEDLINE=21156231; PubMed=11258796;				
RA	Hayashi T., Makino K., Ohnishi M., Kurokawa K., Ishii K., Yokoyama K.,				
RA	Han C.-G., Ohtsubo E., Nakayama K., Murata T., Tanaka M., Tobe T.,				
RA	Iida T., Takami H., Honda T., Sasakawa C., Ogasawara N., Yasunaga T.,				
RA	Kuhara S., Shiba T., Hattori M., Shingawa H.;				
RT	"Complete genome sequence of enterohaemorrhagic Escherichia coli				
RT	O157:H7 and genomic comparison with a laboratory strain K-12."				
RL	DNA Res. 8:11-22 (2001).				

SEQUENCE FROM N.A.  
MEDLINE=86223772; PubMed=2423502;  
Weber P.C., Palchaudhuri S.;  
"Incompatibility repressor in a RepA-like replicon of the IncFI  
plasmid ColV2-K94";  
J. Bacteriol. 166:1106-1112(1986).  
CC -!- FUNCTION: THIS PROTEIN IS ESSENTIAL FOR PLASMID REPLICATION; IT  
CC IS INVOLVED IN COPY CONTROL FUNCTIONS.  
CC -!- MISCELLANEOUS: THIS REPLICATION PROTEIN IS COMPATIBLE WITH THE  
CC REPA REPLICON FOR THE INCFI R PLASMIDS.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
DR EMBL; M13472; AAA23195.1; -;  
KW Plasmid; DNA replication; Plasmid copy control.  
FT NON TER 77  
SQ SEQUENCE 77 AA; 8941 MW; 44784ECC89D965E9 CRC64;  
  
Query Match 30.8%; Score 28; DB 1; Length 77;  
Best Local Similarity 75.0%; Pred. No. 4.4e+02;  
Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;  
  
QY 1 RRVYDAL 8  
DB 68 RRRADAL 75  
||| |||  
-----  
RESULT 24  
VI05\_VACCV STANDARD; PRT; 79 AA.  
AC P20500;  
DT 01-FEB-1991 (Rel. 17, Created)  
DT 01-FEB-1991 (Rel. 17, Last sequence update)  
DT 16-OCT-2001 (Rel. 40, Last annotation update)  
DE Protein I5.  
GN I5L.  
OS Vaccinia virus (strain Copenhagen).  
OC Viruses; dsDNA viruses, no RNA stage; Poxviridae; Chordopoxvirinae;  
OC Orthopoxvirus.  
OX NCBI\_TaxID=10249;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=91021027; PubMed=2219722;  
RA Goebel S.J., Johnson G.P., Perkus M.E., Davis S.W., Winslow J.P.,  
RA Paolletti E.;  
RT "The complete DNA sequence of vaccinia virus";  
RL Virology 179:247-266(1990).  
RN [2]  
RP COMPLETE GENOME.  
RA Goebel S.J., Johnson G.P., Perkus M.E., Davis S.W., Winslow J.P.,  
RA Paolletti E.;  
RT "Appendix to 'The complete DNA sequence of vaccinia virus'";  
RL Virology 179:517-563(1990).  
RN [3]  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
DR EMBL; M35027; AAA48061.1; -;  
DR PIR; A42511; A42511.  
DR InterPro; IPR006803; Pox\_I5.  
DR Pfam; PF04713; Pox\_I5; 1.  
KW Late protein.

SQ SEQUENCE 79 AA; 8762 MW; 76F4826B7009DFAF CRC64;  
  
Query Match 30.8%; Score 28; DB 1; Length 79;  
Best Local Similarity 54.5%; Pred. No. 4.5e+02;  
Matches 6; Conservative 2; Mismatches 3; Indels 0; Gaps 0;  
  
QY 6 DALNVLAMNMI 16  
DB 3 DAITVLTAIGI 13  
||| ||| |||  
-----  
RESULT 25  
VI05\_VACCV STANDARD; PRT; 79 AA.  
AC P12924;  
DT 01-OCT-1989 (Rel. 12, Created)  
DT 01-OCT-1989 (Rel. 12, Last sequence update)  
DT 01-FEB-1991 (Rel. 17, Last annotation update)  
DE Protein I5.  
GN I5L.  
OS Vaccinia virus (strain WR).  
OC Viruses; dsDNA viruses, no RNA stage; Poxviridae; Chordopoxvirinae;  
OC Orthopoxvirus.  
OX NCBI\_TaxID=10254;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=88215015; PubMed=2835495;  
RA Schmitt J.P.C., Stunnenberg H.G.;  
RT "Sequence and transcriptional analysis of the vaccinia virus HindIII  
RT I fragment";  
RL J. Virol. 62:1889-1897(1988).  
RN [2]  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
DR EMBL; J03399; AAB59807.1; -;  
DR PIR; E29889; WZVZ15.  
DR InterPro; IPR006803; Pox\_I5.  
DR Pfam; PF04713; Pox\_I5; 1.  
KW Late protein.  
SQ SEQUENCE 79 AA; 8744 MW; 37F4826B71CF04C3 CRC64;  
  
Query Match 30.8%; Score 28; DB 1; Length 79;  
Best Local Similarity 54.5%; Pred. No. 4.5e+02;  
Matches 6; Conservative 2; Mismatches 3; Indels 0; Gaps 0;  
  
QY 6 DALNVLAMNMI 16  
DB 3 DAITVLTAIGI 13  
||| ||| |||  
-----  
RESULT 26  
VI05\_VARV STANDARD; PRT; 79 AA.  
AC P33001;  
DT 01-OCT-1993 (Rel. 27, Created)  
DT 01-OCT-1993 (Rel. 27, Last sequence update)  
DT 01-FEB-1996 (Rel. 33, Last annotation update)  
DE Protein I5.  
GN I5L OR K5L.  
OS Variola virus.  
OC Viruses; dsDNA viruses, no RNA stage; Poxviridae; Chordopoxvirinae;  
OC Orthopoxvirus.  
OX NCBI\_TaxID=10255;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX STRAIN=India-1967 / Isolate Ind3;  
RX MEDLINE=94152154; PubMed=8109150;



```

OC
CX
COX
NCBI_TaxID=727;
[1]
SEQUENCE FROM N.A.
STRAIN=rd / KW20 / ATCC 51907;
MEDLINE=95350630; PubMed=7542800;
Fleischmann R.D., Adams M.D., White O., Clayton R.A., Kirkness E.F.,
Kerlavage A.R., Bult C.J., Tomb J.-F., Dougherty B.A., Merrick J.M.,
McKenney K., Sutton G., Fitzhugh W., Fields C.A., Gocayne J.D.,
Scott J.D., Shirley R., Liu L.-I., Glodek A., Kelley J.M.,
Weidman J.F., Phillips C.A., Spriggs T., Hedblom E., Cotton M.D.,
Utterback T.R., Hanna M.C., Nguyen D.T., Saudek D.M., Brandon R.C.,
Fine L.D., Fritchman J.D., Fuhrmann J.L., Geoghegan N.S.M.,
Gnehm C.L., McDonald L.A., Small K.V., Fraser C.M., Smith H.O.,
Venter J.C.;
"Whole-genome random sequencing and assembly of Haemophilus influenzae
rd.";
NCBI_TaxID=727;
[2]
SEQUENCE 269:496-512(1995).
SEQUENCE OF 76-88.
MEDLINE=20137488; PubMed=10675023;
Langen H., Takacs B., Evers S., Berndt P., Lahm H.W., Wipf B.,
Gray C., Fountoulakis M.;
"Two-dimensional map of the proteome of Haemophilus influenzae.";
Electrophoresis 21:411-429(2000).
-!- SIMILARITY: BELONGS TO THE VAPD FAMILY.
-----
This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announcement/
or send an email to license@isb-sib.ch).
-----
EMBL; U32728; AAC22108.1; -.
FIR; C64069; C64069.
TIGR; HI0450; -.
Pfam; PF04605; Vapd_N; 1.
Virulence; Complete proteome.
SEQUENCE 91 AA; 10543 MW; 70B23CDBE28E083E7 CRC64;
-----
Query Match 30.8%; Score 28; DB 1; Length 91;
Best Local Similarity 35.7%; Pred. No. 5.3e+02;
Matches 5; Conservative 3; Mismatches 6; Indels 0; Gaps 0;

QY 6 DALNLVLMAMNIISK 19
DB 50 DMANLFQAMNALKQ 63
+ + + + +
+ + + + +

RESULT 28
RT21 HUMAN STANDARD; PRT; 87 AA.
ID RT21 HUMAN STANDARD; PRT; 87 AA.
AC P82921; Q9BST6;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE Mitochondrial 28S ribosomal protein S21 (MRP-S21) (MDS016).
GN MRPS21 OR RPWS21.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
NCBI_TaxID=9606;
[1]
SEQUENCE FROM N.A.
Huang C., Olan B., Tu Y., Gu W., Wang Y., Han Z., Chen Z.;
"Novel genes expressed in hematopoietic stem/progenitor cells from
myelodysplastic syndromes patient.";
RL Submitted (SEP-1999) to the EMBL/GenBank/DBJ databases.
[2]
SEQUENCE FROM N.A.
MEDLINE=21413863; PubMed=11402041;
RX

```





```

RESULT 30
RS15_RICCN STANDARD; PRT; 91 AA.
AC Q92HV6;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE 30S ribosomal protein S15.
GN RPSO OR RC0864.
OS Rickettsia conorii.
OC Bacteria; Proteobacteria; Alphaproteobacteria; Rickettsiales;
OC Rickettsiaceae; Rickettsiae; Rickettsia.
OX NCBI_TaxID=781;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Malish 7;
RX MEDLINE=21442074; PubMed=1157893;
RA Ogata H., Audic S., Renesto-Audiffren P., Fournier P.-E., Barbe V.,
RA Samson D., Roux V., Cossart P., Weissenbach J., Claverie J.-M.,
RA Raoult D.;
RT "Mechanisms of evolution in Rickettsia conorii and R. prowazekii.";
RL Science 293:2093-2098(2001).
CC -1- FUNCTION: THIS PROTEIN IS ONE OF THE 16S RIBOSOMAL RNA BINDING
CC PROTEINS (BY SIMILARITY).
CC -1- SIMILARITY: BELONGS TO THE S15P FAMILY OF RIBOSOMAL PROTEINS.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; AE008625; AL003202.1; -.
DR PIR; H97782; H97782.
DR InterPro; IPR000589; Ribosomal_S15.
DR InterPro; IPR005290; RS15_bact_1.
DR Pfam; PF00312; Ribosomal_S15.
DR ProDom; PD157043; RS15_bact; 1.
DR TIGRFAMs; TIGR00952; S15_bact; 1.
DR PROSITE; PS00362; RIBOSOMAL_S15; 1.
DR Ribosomal protein; rRNA-binding; Complete proteome.
KW Ribosomal protein; rRNA-binding; Complete proteome.
SQ SEQUENCE 91 AA; 10594 MW; 5C6D4C302FA49114 CRC64;

Query Match 30.2%; Score 27.5; DB 1; Length 91;
Best Local Similarity 36.4%; Pred. No. 6.5e+02;
Matches 8; Conservative 6; Mismatches 5; Indels 3; Gaps 1;

QY 1 RRRVYDAL---NVLMAMNIISK 19
DB 63 RRLNVIKKNVSKYLDLSK 84

RESULT 31
UC27_MAIZE STANDARD; PRT; 15 AA.
AC P06633;
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE Unknown protein from 2D-page of etiolated coleoptile (Spot 688)
DE (Fragment).
OS Zea mays (Maize).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC PACCAD clade; Panicoideae; Andropogoneae; Zea.
OX NCBI_TaxID=4577;
RN [1]
RP SEQUENCE.
RC TISSUE=Coleoptile;
RA Touzet P., Riccardi F., Morin C., Damerval C., Huet J.-C.,
RA Pernollet J.-C., Zivy M., de Vienne D.;

Query Match 29.7%; Score 27; DB 1; Length 58;
Best Local Similarity 41.7%; Pred. No. 4.8e+02;
Matches 5; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

QY 6 DALNVLMMNII 17
DB 23 DPLNLSQAQNL 34

RESULT 33
SECE_TREPA STANDARD; PRT; 59 AA.
ID SECE_TREPA
AC O81263;
DT 15-DEC-1998 (Rel. 37, Created)
DT 15-DEC-1998 (Rel. 37, Last sequence update)

```

```

RT "The maize two dimensional gel protein database: towards an integrated
RT genome analysis program.";
RL Theor. Appl. Genet. 93:997-1005(1996).
CC -1- MISCELLANEOUS: ON THE 2D-GEL THE DETERMINED PI OF THIS UNKNOWN
CC PROTEIN IS: 6.4, ITS MW IS: 48.4 kDa.
CC -1- SIMILARITY: TO XENOPUS HISTONE-BINDING PROTEIN N1/N2 AND RABBIT
CC AND HUMAN NUCLEAR AUTOANTIGENIC SPERM PROTEIN.
DR Maize-2DPAGE; P80633; COLEOPTILE.
DR MaizeDB; 123958; -.
FT NON_TER 1 1
FT NON_TER 15 15
SQ SEQUENCE 15 AA; 1853 MW; CA0E12A5DAED8DC7 CRC64;

Query Match 29.7%; Score 27; DB 1; Length 15;
Best Local Similarity 62.5%; Pred. No. 1.1e+02;
Matches 5; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 RRRVYDAL 8
DB 6 RDQVYDAM 13

RESULT 32
YC18_PORPU STANDARD; PRT; 58 AA.
ID YC18_PORPU
AC P51366;
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 15-DEC-1998 (Rel. 37, Last annotation update)
DE Hypothetical 6.9 kDa protein ycf18 (ORF58).
GN YCF18.
OS Porphyra purpurea.
OC Chlorophyta.
OC Eukaryota; Rhodophyta; Bangiophyceae; Bangiales; Porphyra.
OX NCBI_TaxID=2787;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Avonport;
RA Reith M.E., Munnholland J.;
RT "Complete nucleotide sequence of the Porphyra purpurea chloroplast
RT genome".
RL Plant Mol. Biol. Rep. 13:333-335(1995).
CC -1- SIMILARITY: SOME, TO SYNECHOCOCCUS PCC 7942 NBLA.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; U38804; AAC08252.1; -.
DR PIR; S73287; S73287.
DR Pfam; PF04485; nbla; 1.
DR Chloroplast; Hypothetical protein.
SQ SEQUENCE 58 AA; 6923 MW; 304D4D9EDD0D2371 CRC64;

Query Match 29.7%; Score 27; DB 1; Length 58;
Best Local Similarity 41.7%; Pred. No. 4.8e+02;
Matches 5; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

QY 6 DALNVLMMNII 17
DB 23 DPLNLSQAQNL 34

RESULT 33
SECE_TREPA STANDARD; PRT; 59 AA.
ID SECE_TREPA
AC O81263;
DT 15-DEC-1998 (Rel. 37, Created)
DT 15-DEC-1998 (Rel. 37, Last sequence update)

```



```

Matches      4;  Conservative      2;  Mismatches      0;  Indels      0;  Gaps      0;

QY      1 RRVYD 6
Db      |||:
      2 RRIYE 7

RESULT 36
DLTC STAXY STANDARD; PRT; 78 AA.
AC Q9X2N6;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE D-alanine--poly(phosphoribitol)ligase subunit 2 (EC 6.1.1.13) (D-
DLTC alanyl carrier protein) (DCP).
OS Staphylococcus xylosum.
OC Bacteria; Firmicutes; Bacillales; Staphylococcus.
OX NCBI_TaxID=1288;
RN SEQUENCE FROM N.A.
RC STRAIN=DSM 20267 / Isolate C2A;
RX MEDLINE=99185055; PubMed=10085071;
RA Peschel A., Otto M., Jack R.W., Kalbacher H., Jung G., Gotz F.;
RT "Inactivation of the dlt operon in Staphylococcus aureus confers
RT sensitivity to defensins, protegrins, and other antimicrobial
RT peptides.;"
RL J. Biol. Chem. 274:8405-8410(1999).
CC -I- FUNCTION: Involved in the biosynthesis of D-alanyl-lipoteichoic
CC acid (LTA). Activated D-alanyl-Dcp donates its D-alanyl
CC substituent to membrane-associated LTA (By similarity).
CC -I- CATALYTIC ACTIVITY: ATP + D-alanine + poly(ribitol phosphate) =
CC AMP + diphosphate + O-D-alanyl-poly(ribitol phosphate).
CC -I- PATHWAY: D-alanyl-lipoteichoic acid biosynthesis.
CC -I- PTM: 4'-phosphopantetheine is transferred from CoA to a specific
CC serine of apo-DCP (By similarity).
CC -I- SIMILARITY: Contains 1 acyl carrier domain.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; AF032440; AAD01944.1; -
CC HAMAP; MF 00565; -; 1.
CC InterPro; IPR003230; D-ala carrier.
CC InterPro; IPR006163; Pp bind.
CC ProDom; PD015103; D-ala carrier; 1.
CC PROSITE; PS00075; ACP_DOMAIN; FALSE_NEG.
CC KW Ligase; Cell wall; Phosphopantetheine.
CC FT BINDING 36 36 PHOSPHOPANTETHEINE (PROBABLE).
CC SQ SEQUENCE 78 AA; 9155 MW; C64B76620E7074C2 CRC64;

Query Match 29.7%; Score 27; DB 1; Length 78;
Best Local Similarity 31.6%; Pred. No. 6.7e+02;
Matches 6; Conservative 4; Mismatches 9; Indels 0; Gaps 0;

QY      1 RRVYDALNVLMMNNIISK 19
Db      |||:
      4 REQVLDTLLEVAENNVK 22

RESULT 37
YNIL FRAAL STANDARD; PRT; 82 AA.
AC P46041;
DT 01-NOV-1995 (Rel. 32, Created)
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)

Hypothetical 9.1 kDa protein in nifX-nifW intergenic region (ORF1).
Frankia alni.
Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
Frankineae; Frankiaceae; Frankia.
NCBI_TaxID=1859;
RN SEQUENCE FROM N.A.
RC STRAIN=Cp11;
RX MEDLINE=95369734; PubMed=7642138;
RA Harriott O.T., Hosted T.J., Benson D.R.;
RT "Sequences of nifX, nifW, nifZ, nifB and two ORF in the Frankia
RT nitrogen fixation gene cluster.;"
RL Gene 161:63-67(1995).
CC -I- SIMILARITY: TO SIMILAR PROTEINS IN OTHER NITROGEN-FIXING BACTERIA.
CC THIS PROTEIN IS GENERALLY FOUND IN THE NIFX-NIFW INTERGENIC REGION
CC OR IN THE FIXX 3'REGION
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; L29299; AAC82972.1; -
CC PIR; T09234; T09234.
CC Pfam; PF05082; DUF683; 1.
CC KW Hypothetical protein; Nitrogen fixation.
CC SQ SEQUENCE 82 AA; 9081 MW; AFBD06827B4322C CRC64;

Query Match 29.7%; Score 27; DB 1; Length 82;
Best Local Similarity 83.3%; Pred. No. 7.1e+02;
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY      2 RRVYDA 7
Db      |||
      60 RRTYDA 65

RESULT 38
RL31 THEAC STANDARD; PRT; 89 AA.
AC Q9HMI7;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE 50S ribosomal protein L31e.
GN RPL31E OR TA0054.
OS Thermoplasma acidophilum.
OC Archaea; Euryarchaeota; Thermoplasmata; Thermoplasmatales;
OC Thermoplasmataceae; Thermoplasma.
OX NCBI_TaxID=2303;
RN SEQUENCE FROM N.A.
RC STRAIN=DSM 1728;
RX MEDLINE=20479972; PubMed=11029001;
RA Ruepp A., Graml W., Santos-Martinez M.-L., Koretke K.K., Volker C.,
RA Mewes H.-W., Frishman D., Stocker S., Lupas A.N., Baumeister W.;
RT "The genome sequence of the thermoacidophilic scavenger Thermoplasma
RT acidophilum.;"
RL Nature 407:508-513(2000).
CC -I- SIMILARITY: BELONGS TO THE L31E FAMILY OF RIBOSOMAL PROTEINS.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; AL445063; CAC11202.1; ALT_INIT.

```

```

DR HAMAP; MF 00410; -. 1.
DR InterPro; IPR000054; Ribosomal_L31e.
DR Pfam; PF01198; Ribosomal_L31e; 1.
DR PROSITE; PS01144; RIBOSOMAL_L31e; FALSE_NEG.
KW Ribosomal protein; Complete proteome.
SQ SEQUENCE 89 AA; 10173 MW; 367E5A1117EAC1AE CRC64;

Query Match 29.7%; Score 27; DB 1; Length 89;
Best Local Similarity 36.8%; Pred. No. 7.7e+02;
Matches 7; Conservative 5; Mismatches 5; Indels 2; Gaps 1;

Qy 1 RRVYDALNVLMMNIISK 19
Db 25 KRRADTAVSIL--RNFVSK 41

RESULT 39
BAF2_HUMAN STANDARD; PRT; 90 AA.
AC Q9H503;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Hypothetical BAF-like protein C20orf179.
GN C20ORF179.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=21638749; PubMed=11780052;
RA Deloukas P., Matthews L.H., Ashurst J., Burton J., Gilbert J.G.R.,
RA Jones M., Stavrides G., Almeida J.P., Babbage A.K., Baguley C.L.,
RA Bailey J., Barlow K.P., Bates K.N., Beard L.M., Beare D.M.,
RA Beasley O.P., Bird C.P., Blakey S.E., Bridgeman A.M., Brown A.J.,
RA Buck D., Burrill W.D., Butler A.P., Carder C., Carter N.P.,
RA Chapman J.C., Clamp M., Clark G., Clark L.N., Clark S.Y., Clee C.M.,
RA Clegg S., Cobley V.E., Collier R.E., Connor R.E., Corby N.R.,
RA Coulson A., Coville G.J., Deadman R., Dhami P.D., Dunn M.,
RA Ellington A.G., Frankland J.A., Fraser A., French L., Garner P.,
RA Grahnam D.V., Griffiths C., Griffiths M.N.D., Gwilliam R., Hall R.E.,
RA Hammond S., Harley J.L., Heath P.D., Ho S., Holden J.L., Howden P.J.,
RA Huckle E., Hunt A.R., Hunt S.E., Jekosch K., Johnson C.M., Johnson D.,
RA Kay M.P., Kimberley A.M., King A., Knights A., Laird G.K., Lawlor S.,
RA Leivaeslaih M.H., Leversha M.A., Lloyd C., Lloyd D.M., Lovell J.D.,
RA Marsh V.L., Martin S.L., McConnachie L.J., McClay K., McMurray A.A.,
RA Milne S.A., Mistry D., Moore M.J.F., Mullikin J.C., Nickerson T.,
RA Oliver K., Parker A., Patel R., Pearce T.A.V., Peck A.I.,
RA Phillimore B.J.C.T., Prathalingam S.R., Plumb R.W., Ramsay H.,
RA Rice C.M., Ross M.T., Scott C.E., Sehra H.K., Showlkeen R., Sims S.,
RA Skuane C.D., Smith M.L., Soderlund C., Steward C.A., Suleston J.E.,
RA Swann R.M., Sycamore N., Taylor R., Tee L., Thomas D.W., Thorpe A.,
RA Tracey A., Tromans A.C., Vaudin M., Wall M., Wallis J.M.,
RA Whitehead S.L., Whittaker P., Willey D.L., Williams L., Williams S.A.,
RA Wilming L., Wray P.W., Hubbard T., Durbin R.M., Bentley D.R., Beck S.,
RA Rogers J.;
RT "The DNA sequence and comparative analysis of human chromosome 20.";
RL Nature 414:865-871(2001).
CC -1- FUNCTION: DNA-binding protein (Potential).
CC -1- SUBCELLULAR LOCATION: Nuclear (Potential).
CC -1- SIMILARITY: BELONGS TO THE BAF FAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC -----
CC EMBL; AL160071; CAC15540.1; -.
CC HSSP; 075531; 2BZX.
CC
DR HAMAP; HGNC:16172; C20orf179.
DR InterPro; IPR004122; BAF_prot.
DR Pfam; PF02961; BAF; 1.
DR PROSITE; PS01144; RIBOSOMAL_L31e; FALSE_NEG.
KW Hypothetical protein; DNA-binding; Nuclear protein; Polymorphism.
FT VARIANT 78 78 /FTID=VAR_013693.
SQ SEQUENCE 90 AA; 10309 MW; C7FAB57610ADEF87 CRC64;

Query Match 29.7%; Score 27; DB 1; Length 90;
Best Local Similarity 28.6%; Pred. No. 7.8e+02;
Matches 4; Conservative 7; Mismatches 3; Indels 0; Gaps 0;

Qy 6 DALNVLMMNIISK 19
Db 25 DGISHELAINLVTK 38

RESULT 40
ILG1_CAEEL STANDARD; PRT; 91 AA.
AC Q18060;
DT 15-JUL-1998 (Rel. 36, Created)
DT 15-JUL-1998 (Rel. 36, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Probable insulin-like peptide gamma-type 1 precursor (Ceinsulin-3).
GN INS-11 OR C17C3.4.
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
OC Rhabditidae; Peloderinae; Caenorhabditis.
OX NCBI_TaxID=6239;
RN [1]
RP SEQUENCE FROM N.A.
RX Kawanoto T.;
RT "mRNA for a putative insulin-like peptide of Caenorhabditis
RT elegans.";
RL Submitted (SEP-1999) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RX STRAIN=Bristol N2;
RA Du Z.;
RL Submitted (DEC-1995) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SIMILARITY TO INSULIN.
RX MEDLINE=98217375; PubMed=9548970;
RA Duret L., Guex N., Peitsch M.C., Bairoch A.;
RT "New insulin-like proteins with atypical disulfide bond pattern
RT characterized in Caenorhabditis elegans by comparative sequence
RT analysis and homology modeling.";
RL Genome Res. 8:348-353(1998).
CC -1- SUBCELLULAR LOCATION: Secreted (Potential).
CC -1- SIMILARITY: BELONGS TO THE INSULIN/IGF/RELAXIN FAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC -----
CC EMBL; AB032258; BRA84470.1; -.
CC EMBL; U41279; AAK31418.1; -.
CC PIR; T37327; T37327.
CC WormPep; C17C3.4; CE04024.
DR InterPro; IPR004825; Ins/IGF/relax.
DR PROSITE; PS00262; INSULIN; 1.
KW Insulin family; Signal.
FT SIGNAL 1 26
FT CHAIN 27 91
FT PROBABLE INSULIN-LIKE PEPTIDE GAMMA-TYPE
FT 1.
FT PEPTIDE 34 58
FT PEPTIDE 61 91
FT DISULFID 37 66

```

```
FT DISULFID 49 79 POTENTIAL.
FT DISULFID 65 70 POTENTIAL.
SQ SEQUENCE 91 AA; 10173 MW; 22BF958BF759F254 CRC64;

Query Match 29.7%; Score 27; DB 1; Length 91;
Best Local Similarity 44.4%; Pred. No. 7.9e+02;
Matches 4; Conservative 4; Mismatches 1; Indels 0; Gaps 0;

QY 3 RYVDALNV 11
Db 40 KIFKALNV 48

RESULT 41
DBH_BUCAI STANDARD; PRT; 92 AA.
AC P57144;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE DNA-binding protein HU.
GN HUP OR BU032.
OS Buchnera aphidicola (subsp. Acyrthosiphon pisum) (Acyrthosiphon pisum
OC symbiotic bacterium).
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Buchnera.
OX NCBI_TaxID=118099;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Tokyo 1998;
RX MEDLINE=20445173; PubMed=10993077;
RA Shigenobu S., Watanabe H., Hattori M., Sakaki Y., Ishikawa H.;
RT "Genome sequence of the endocellular bacterial symbiont of aphids
RT Buchnera sp. APS.";
RL Nature 407:81-86(2000).
CC -!- FUNCTION: THIS PROTEIN BELONGS TO THE HISTONE LIKE FAMILY OF
CC PROKARYOTIC DNA-BINDING PROTEINS WHICH ARE CAPABLE OF WRAPPING
CC DNA TO STABILIZE IT, AND PREVENT ITS DENATURATION UNDER EXTREME
CC ENVIRONMENTAL CONDITIONS (BY SIMILARITY).
CC -!- SUBUNIT: Homodimer (By similarity).
CC -!- SIMILARITY: BELONGS TO THE BACTERIAL HISTONE-LIKE PROTEIN FAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC -----
CC EMBL; AP001118; BAB12759.1; -.
CC HSSP; P36206; 188Z.
CC InterPro; IPR000119; Bac DNABind.
CC Pfam; PF00216; Bac DNA binding; 1.
CC ProDom; PD000945; Bac DNABind; 1.
CC SMART; SM00411; BHL; 1.
CC PROSITE; PS00045; HISTONE LIKE; 1.
CC DNA-binding; DNA condensation; Complete proteome.
CC SEQUENCE 92 AA; 10044 MW; 998475DABE888118 CRC64;

Query Match 29.7%; Score 27; DB 1; Length 92;
Best Local Similarity 41.7%; Pred. No. 8e+02;
Matches 5; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

QY 8 LNVLMANNIISK 19
Db 1 MNKTOLINVISK 12

RESULT 42
RPC1_BPP22 STANDARD; PRT; 92 AA.
ID RPC1_BPP22
AC P03041;

FT DISULFID 49 79 POTENTIAL.
FT DISULFID 65 70 POTENTIAL.
SQ SEQUENCE 91 AA; 10173 MW; 22BF958BF759F254 CRC64;

Query Match 29.7%; Score 27; DB 1; Length 91;
Best Local Similarity 44.4%; Pred. No. 7.9e+02;
Matches 4; Conservative 4; Mismatches 1; Indels 0; Gaps 0;

QY 3 RYVDALNV 11
Db 40 KIFKALNV 48

RESULT 43
VAPD_ACTAC STANDARD; PRT; 95 AA.
ID VAPD_ACTAC
AC Q52243;
DT 15-JUL-1999 (Rel. 38, Created)
```

```
DT 21-JUL-1986 (Rel. 01, Created)
DT 21-JUL-1986 (Rel. 01, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Transcriptional activator protein C1.
GN C1.
OS Bacteriophage P22.
OC Viruses; dsDNA viruses, no RNA stage; Caudovirales; Podoviridae;
OC P22-like viruses.
OX NCBI_TaxID=10754;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=85155495; PubMed=6241581;
RA Backhaus H., Petri J.B.;
RT "Sequence analysis of a region from the early right operon in phage
RT P22 including the replication genes 18 and 12.";
RL Gene 32:289-303(1984).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=86159692; PubMed=3954988;
RA Poteete A.R.;
RT "Bacteriophage P22 Cro protein: sequence, purification, and
RT properties.";
RL Biochemistry 25:251-256(1986).
RN [3]
RP SEQUENCE FROM N.A.
RA Kropinski A.M.B., VanderByl C.S.;
RT "The completed sequence of genome of salmonella phage P22.";
RL Submitted (DEC-1999) to the EMBL/GenBank/DBJ databases.
RN [4]
RP SEQUENCE OF 1-84, AND CHARACTERIZATION.
RX MEDLINE=92332555; PubMed=1385814;
RA Ho Y.S., Pfarr D., Strickler J., Rosenberg M.;
RT "Characterization of the transcription activator protein C1 of
RT bacteriophage P22.";
RL J. Biol. Chem. 267:14388-14397(1992).
CC -!- FUNCTION: BINDS TO TWO PROMOTERS, P(RE) AND PA23 AND ACTIVATE
CC TRANSCRIPTION FROM THESE PROMOTERS.
CC -!- SUBUNIT: Homotetramer.
CC -!- SIMILARITY: THIS PROTEIN IS RELATED TO THE REGULATORY PROTEIN
CC CII OF BACTERIOPHAGES LAMBDA AND 434.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC -----
CC EMBL; M10074; AAA32274.1; -.
CC EMBL; M12584; AAA32269.1; -.
CC EMBL; AF217253; AAF75026.1; -.
CC PIR; A91518; Z1BPC2.
CC Pfam; PF05269; Phage CII; 1.
CC Transcription regulation; DNA-binding; Activator.
CC DNA_BIND 26 45 H-T-H MOTIF (PROBABLE).
CC CONFLICT 83 83 P -> D (IN REF. 4)
CC SEQUENCE 92 AA; 10211 MW; D479C6D90085CEB6 CRC64;

Query Match 29.7%; Score 27; DB 1; Length 92;
Best Local Similarity 50.0%; Pred. No. 8e+02;
Matches 5; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 1 RRRVYDALNV 10
Db 26 QRKVADALGI 35

RESULT 43
VAPD_ACTAC STANDARD; PRT; 95 AA.
ID VAPD_ACTAC
AC Q52243;
DT 15-JUL-1999 (Rel. 38, Created)
```

DT 15-JUL-1999 (Rel. 38, Last sequence update)  
 DT 15-JUL-1999 (Rel. 38, Last annotation update)  
 DE Virulence-associated protein D homolog.  
 OS *Actinobacillus actinomycetemcomitans* (Haemophilus  
 OG Plasmid pVT736-1.  
 OC Bacteria; Proteobacteria; Gammaproteobacteria; Pasteurellales;  
 OC Pasteurellaceae; *Actinobacillus*.  
 OX NCBI\_TaxID=714;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=95362676; PubMed=7543479;  
 RA Galli D.M., Leblanc D.J.;  
 RT "Transcriptional analysis of rolling circle replicating plasmid  
 RL pVT736-1: evidence for replication control by antisense RNA.";  
 RL J. Bacteriol. 177:4474-4480(1995).  
 CC -1- SIMILARITY: BELONGS TO THE VAPD FAMILY.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 CC EMBL; L24000; AAC37126.1; -;  
 DR PFam; PF04605; VapD\_N; 1.  
 KW Plasmid; Virulence.  
 SQ SEQUENCE 95 AA; 11190 MW; 3424348E815BF62A CRC64;  
 Query Match 29.7%; Score 27; DB 1; Length 95;  
 Best Local Similarity 62.5%; Pred. No. 8.3e+02;  
 Matches 5; Conservative 2; Mismatches 1; Indels 0; Gaps 0;  
 QY 4 VYDALNVL 11  
 DB 57 VYKAINLL 64  
 RESULT 44  
 C553\_HELPJ STANDARD; PRT; 96 AA.  
 ID Q9ZJ29;  
 DT 28-FEB-2003 (Rel. 41, Created)  
 DT 28-FEB-2003 (Rel. 41, Last sequence update)  
 DT 28-FEB-2003 (Rel. 41, Last annotation update)  
 DE Cytochrome C-553 precursor (C553).  
 GN JHP1148.  
 OS *Helicobacter pylori* J99 (Campylobacter pylori J99).  
 OC Bacteria; Proteobacteria; Epsilonproteobacteria; Campylobacterales;  
 OC Helicobacteraceae; *Helicobacter*.  
 OX NCBI\_TaxID=85963;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=99120557; PubMed=9923682;  
 RA Alm R.A., Ling L.-S.L., Moir D.T., King B.L., Brown E.D., Doig P.C.,  
 RA Smith D.R., Noonan B., Guild B.C., deJonge B.L., Carmel G.,  
 RA Tummino P.J., Caruso A., Uria-Nickelsen M., Mills D.M., Ives C.,  
 RA Gibson R., Merberg D., Mills S.D., Jiang Q., Taylor D.E., Vovis G.F.,  
 RA Trust T.J.;  
 RT "Genomic sequence comparison of two unrelated isolates of the human  
 RL gastric pathogen *Helicobacter pylori*";  
 RL Nature 387:176-180(1999).  
 CC -1- FUNCTION: Natural electron acceptor for a formate dehydrogenase.  
 CC -1- SUBCELLULAR LOCATION: Periplasmic.  
 CC -1- PTM: BINDS ONE HEME GROUP PER MOLECULE.  
 CC -1- SIMILARITY: BELONGS TO THE CLASS I CYTOCHROME C FAMILY.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way

CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 CC EMBL; AE001542; AAD06721.1; -;  
 DR PIR; F71843; F71843.  
 DR HSSP; P04032; 2DVH.  
 DR InterPro; IPR003088; Cyt\_C1.  
 DR InterPro; IPR002329; Cyt\_C1C.  
 DR InterPro; IPR000345; CytC\_heme\_bind.  
 DR Pfam; PF00034; cytochrome\_C\_1.  
 DR PRINTS; PR00605; CYTOCHROME\_C1.  
 DR ProDom; PD004020; Cyt\_C\_bact; 1.  
 DR PROSITE; PS00190; CYTOCHROME\_C; 1.  
 KW Electron transport; Heme; Signal; Periplasmic; Complete proteome.  
 FT SIGNAL 1 19  
 FT CHAIN 20 95 CYTOCHROME C-553.  
 FT BINDING 29 29 HEME (COVALENT).  
 FT BINDING 32 32 HEME (COVALENT).  
 FT METAL 33 33 IRON (HEME AXIAL LIGAND) (BY  
 FT METAL 73 73 IRON (HEME AXIAL LIGAND) (BY  
 FT SIMILARITY).  
 SQ SEQUENCE 96 AA; 10354 MW; 3E607AE5D422AD82 CRC64;  
 Query Match 29.7%; Score 27; DB 1; Length 96;  
 Best Local Similarity 37.5%; Pred. No. 8.4e+02;  
 Matches 6; Conservative 4; Mismatches 6; Indels 0; Gaps 0;  
 QY 2 RRYVDALNVLMMANVII 17  
 DB 2 KKVIVAGVLAFANVL 17  
 RESULT 45  
 NULM\_BALMU STANDARD; PRT; 98 AA.  
 ID P41301;  
 DT 01-FEB-1995 (Rel. 31, Created)  
 DT 01-FEB-1995 (Rel. 31, Last sequence update)  
 DT 01-NOV-1997 (Rel. 35, Last annotation update)  
 DE NADH-ubiquinone oxidoreductase chain 4L (EC 1.6.5.3).  
 GN MTND4L OR ND4L OR NADH4L.  
 OS Balaenoptera musculus (Blue whale).  
 OG Mitochondrion.  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Cetartiodactyla; Cetacea; Mysticeti;  
 OC Balaenopteridae; Balaenoptera.  
 OX NCBI\_TaxID=9771;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=94141932; PubMed=8308901;  
 RA Arnason U., Gullberg A.;  
 RT "Comparison between the complete mtDNA sequences of the blue and the  
 RL fin whale, two species that can hybridize in nature.";  
 RL J. Mol. Evol. 37:312-322(1993).  
 CC -1- CATALYTIC ACTIVITY: NADH + ubiquinone = NAD(+) + ubiquinol.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 CC EMBL; X72204; CAA51003.1; -;  
 DR PIR; S41828; S41828.  
 DR InterPro; IPR001133; Oxidored4L.  
 DR InterPro; IPR003214; Oxidred4L.  
 DR Pfam; PF00420; Oxidored\_q2; 1.  
 DR ProDom; PD000359; Oxidred4L; 1.  
 KW Oxidoreductase; NAD; Ubiquinone; Mitochondrion.

SQ SEQUENCE 98 AA; 10747 MW; 9F770651FE6SEED1B CRC64;  
Query Match 29.7%; Score 27; DB 1; Length 98;  
Best/Local Similarity 55.6%; Pred. No. 8.6e+02;  
Matches 5; Conservative 3; Mismatches 1; Indels 0; Gaps 0;  
QY 8 LNVLMAMNI 16  
:||||| :  
Db 6 MNVLMAFSM 14

Search completed: February 17, 2004, 10:57:02  
Job time : 7.0198 secs



GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: February 17, 2004, 10:50:12 ; Search time 26.901 Seconds  
(without alignments)  
182.261 Million cell updates/sec

Title: US-09-900-147-3

Perfect score: 91  
Sequence: 1 RRRVYDALNVLMMNNIIISK 19

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 830525 seqs, 258052604 residues

Total number of hits satisfying chosen parameters: 146963

Minimum DB seq length: 0  
Maximum DB seq length: 100

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : SPTRMBL\_23.\*

- 1: sp\_archaea.\*
- 2: sp\_bacteria.\*
- 3: sp\_fungi.\*
- 4: sp\_human.\*
- 5: sp\_invertebrate.\*
- 6: sp\_mammal.\*
- 7: sp\_mmc.\*
- 8: sp\_organelle.\*
- 9: sp\_phase.\*
- 10: sp\_plant.\*
- 11: sp\_rodent.\*
- 12: sp\_virus.\*
- 13: sp\_vertebrate.\*
- 14: sp\_unclassified.\*
- 15: sp\_virus.\*
- 16: sp\_bacteriaph.\*
- 17: sp\_archaeap.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	40	44.0	58	17 Q8ZV98	Q8ZV98 pyrobaculum
2	39	42.9	80	11 Q99MV9	Q99MV9 mus musculus
3	37	40.7	71	10 Q9LQF4	Q9LQF4 arabidopsis
4	35	38.5	39	16 Q8EAL3	Q8EAL3 shewanella
5	35	38.5	58	16 Q98M32	Q98M32 rhizobium l
6	35	38.5	68	16 Q8DSX0	Q8DSX0 streptococc
7	35	38.5	80	11 Q8C8H4	Q8C8H4 mus musculus
8	35	38.5	83	10 Q42022	Q42022 arabidopsis
9	34	37.4	74	16 Q99VZ5	Q99VZ5 staphylococ
10	34	37.4	74	17 Q9HHQ1	Q9HHQ1 halobacteri
11	34	37.4	77	17 Q28868	Q28868 archaeoglob
12	34	37.4	82	9 Q48472	Q48472 bacterioph
13	34	37.4	99	16 Q8E4Q1	Q8E4Q1 streptococ
14	33	36.3	66	17 Q9HQ82	Q9HQ82 halobacteri
15	33	36.3	66	17 Q979G1	Q979G1 thermoplasm
16	33	36.3	79	13 Q91038	Q91038 gadus morhu

17	33	36.3	83	16 Q8EIV8	Q8EIV8 shewanella
18	33	36.3	92	8 Q63454	Q63454 cupha eryma
19	33	36.3	94	16 Q8XYL3	Q8XYL3 ralstonia s
20	33	36.3	95	4 Q96DE8	Q96DE8 homo sapien
21	32.5	35.7	89	9 Q8HA80	Q8HA80 bacterioph
22	32	35.2	24	5 Q44050	Q44050 trypanosoma
23	32	35.2	53	16 Q8X3F6	Q8X3F6 escherichia
24	32	35.2	55	8 P92499	P92499 apis mellif
25	32	35.2	55	8 P92501	P92501 apis mellif
26	32	35.2	60	17 Q26288	Q26288 methanobact
27	32	35.2	61	2 Q9Z682	Q9Z682 bradyrhizob
28	32	35.2	64	12 Q8JKS1	Q8JKS1 heliothis z
29	32	35.2	66	16 Q99RD1	Q99RD1 staphylococ
30	32	35.2	71	6 Q9GMS9	Q9GMS9 macaca fasc
31	32	35.2	72	10 Q94EX1	Q94EX1 arabidopsis
32	32	35.2	77	8 Q34371	Q34371 dryadula ph
33	32	35.2	78	13 Q57412	Q57412 tetraodon f
34	32	35.2	83	16 Q9K761	Q9K761 bacillus ha
35	32	35.2	89	2 Q54418	Q54418 serratia ma
36	32	35.2	94	17 Q9HRH0	Q9HRH0 halobacteri
37	32	35.2	96	12 Q8V6V2	Q8V6V2 halovirus h
38	32	35.2	96	17 Q30267	Q30267 archaeoglob
39	31.5	34.6	52	9 Q38507	Q38507 bacterioph
40	31	34.1	35	16 Q9KAU0	Q9KAU0 bacillus ha
41	31	34.1	49	13 Q9ES47	Q9ES47 bothrops ja
42	31	34.1	62	16 Q8F925	Q8F925 leptospira
43	31	34.1	68	16 Q8YP72	Q8YP72 anabaena sp
44	31	34.1	69	17 Q97AT0	Q97AT0 thermoplasm
45	31	34.1	70	4 Q96IE9	Q96IE9 homo sapien

## ALIGNMENTS

## RESULT 1

Q8ZV98	ID	Q8ZV98	PRELIMINARY;	PRT;	58 AA.
AC	Q8ZV98;				
DT	01-MAR-2002 (TrEMBLrel. 20, Created)				
DT	01-MAR-2002 (TrEMBLrel. 20, Last sequence update)				
DT	01-MAR-2002 (TrEMBLrel. 20, Last annotation update)				
DE	Hypothetical protein PAE2389.				
GN	PAE2389.				
OS	Pyrobaculum aerophilum.				
OC	Archaea; Crenarchaeota; Thermoprotei; Thermoproteales;				
OC	Thermoproteaceae; Pyrobaculum.				
OX	NCBI_TaxID=13773;				
RN	[1]				
RP	SEQUENCE FROM N.A.				
RC	STRAIN=IM2 / ATCC 51768 / DSM 7523;				
RX	PubMed=11792869;				
RA	Fitz-Gibbon S.T., Ladner H., Kim U.-J., Stetter K.O., Simon M.I.,				
RA	Miller J.H.;				
RT	"Genome sequence of the hyperthermophilic crenarchaeon Pyrobaculum				
RT	aerophilum.";				
RL	Proc. Natl. Acad. Sci. U.S.A. 99:984-989(2002).				
DR	EMBL; AE009873; AAL64158.1; -				
KW	Hypothetical protein; Complete proteome.				
SQ	SEQUENCE 58 AA; 6342 MW; 110191142AD83792 CRC64;				

Query Match 44.0%; Score 40; DB 17; Length 58;  
Best Local Similarity 47.1%; Pred. No. 25;  
Matches 8; Conservative 3; Mismatches 6; Indels 0; Gaps 0;

Qy 1 RRRVYDALNVLMMNNII 17

Db 3 RRRYEGNLNPFVAAGLI 19

## RESULT 2

Q99MV9	ID	Q99MV9	PRELIMINARY;	PRT;	80 AA.
AC	Q99MV9;				



```

Qy 1 RRVYDALNVLMANNII 17
    :||: : :||:
Db 10 KRLRTERVDIIAAIL 26

RESULT 6
Q8DSX0 PRELIMINARY; PRT; 68 AA.
AC Q8DSX0
DT 01-MAR-2003 (T-EMBLrel. 23, Created)
DT 01-MAR-2003 (T-EMBLrel. 23, Last sequence update)
DT 01-MAR-2003 (T-EMBLrel. 23, Last annotation update)
DE Hypothetical protein.
GN SMU.1637C.
OS Streptococcus mutans.
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC Streptococcus.
OX NCBI_TaxID=1309;
RN [1]
RC SEQUENCE FROM N.A.
RC STRAIN=UA159 / ATCC 700610 / Serotype C;
RX MEDLINE=22295063; PubMed=12397186;
RA Ajdic D., McShan W.M., McLaughlin R.E., Savic G., Chang J.,
RA Carson M.B., Primeaux C., Tian R., Kenton S., Jia H., Lin S., Qian Y.,
RA Li S., Zhu H., Najjar F., Lai H., White J., Roe B.A., Ferretti J.J.;
RT "Genome sequence of Streptococcus mutans UA159, a cariogenic dental
RT pathogen.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:14434-14439(2002).
DR EMBL; AE014994; AA059277.1; -.
KW Hypothetical protein; Complete proteome.
SQ SEQUENCE 68 AA; 7604 MW; EEF457B026865773 CRC64;

Query Match 38.5%; Score 35; DB 16; Length 68;
Best Local Similarity 53.8%; Pred. No. 2.1e+02;
Matches 7; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

Qy 6 DALNVLMANNIIIS 18
    :||: ||| :||:
Db 2 DALNHLMTNLT 14

RESULT 7
Q8C8H4 PRELIMINARY; PRT; 80 AA.
AC Q8C8H4
DT 01-MAR-2003 (T-EMBLrel. 23, Created)
DT 01-MAR-2003 (T-EMBLrel. 23, Last sequence update)
DT 01-MAR-2003 (T-EMBLrel. 23, Last annotation update)
DE Unknown EST (Fragment).
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RC SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Cerebellum;
RX MEDLINE=22354683; PubMed=12466851;
RA The FANTOM Consortium,
RA "Analysis of the mouse transcriptome based on functional annotation of
RT 60,770 full-length cDNAs.",
RL Nature 420:563-573(2002).
DR EMBL; AK047089; BAC32957.1; -.
FT NON_TER 1
SQ SEQUENCE 80 AA; 9655 MW; D03C342182DC4BD2 CRC64;

Query Match 38.5%; Score 35; DB 11; Length 80;
Best Local Similarity 50.0%; Pred. No. 2.5e+02;
Matches 6; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

Qy 4 VYDALNVLMANN 15
    :||: ||| :||:

```

```

Db 18 IYNALKLFEMMN 29

RESULT 8
Q42022 PRELIMINARY; PRT; 83 AA.
AC Q42022
DT 01-NOV-1996 (T-EMBLrel. 01, Created)
DT 01-NOV-1996 (T-EMBLrel. 01, Last sequence update)
DT 01-OCT-2002 (T-EMBLrel. 22, Last annotation update)
DE Vacuolar ATP synthase 57KD subunit (Fragment).
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eurosids II; Brassicales; Brassicaceae; Arabidopsis.
OX NCBI_TaxID=3702;
RN [1]
RC SEQUENCE FROM N.A.
RC STRAIN=cv. Columbia;
RA Berthomieu P., Guerrier D., Giraudat J.;
RL Submitted (NOV-1992) to the EMBL/GenBank/DBJ databases.
CC -1- SIMILARITY: BELONGS TO THE ATPASE ALPHA/BETA CHAINS FAMILY.
DR EMBL; Z18510; CAA79211.1; -.
DR InterPro; IPR000194; ATPase_a/bcentre.
DR Pfam; PF00006; ATP-synt_ab; 1.
DR PROSITE; PS00152; ATPASE ALPHA BETA; 1.
KW Hydrogen ion transport; Hydrolase; Ion transport; Transport.
FT NON_TER 1
FT NON_TER 83
SQ SEQUENCE 83 AA; 9443 MW; 1B6EBB392EEEC550 CRC64;

Query Match 38.5%; Score 35; DB 10; Length 83;
Best Local Similarity 33.3%; Pred. No. 2.6e+02;
Matches 6; Conservative 7; Mismatches 5; Indels 0; Gaps 0;

Qy 2 RRVYDALNVLMANNIIISK 19
    :||: ||| :||:
Db 38 RQIYPPINVLPSLRMK 55

RESULT 9
Q99VZ5 PRELIMINARY; PRT; 74 AA.
AC Q99VZ5
DT 01-JUN-2001 (T-EMBLrel. 17, Created)
DT 01-JUN-2001 (T-EMBLrel. 17, Last sequence update)
DT 01-OCT-2002 (T-EMBLrel. 22, Last annotation update)
DE Hypothetical protein SAV0618 (Hypothetical protein MW0592).
GN SAV0618 OR SA0575 OR MW0592.
OS Staphylococcus aureus (strain Mu50 / ATCC 700699),
OS Staphylococcus aureus (strain N315), and
OS Staphylococcus aureus (strain MW2).
OC Bacteria; Firmicutes; Bacillales; Staphylococcus.
OX NCBI_TaxID=158878, 158879, 196620;
RN [1]
RC SEQUENCE FROM N.A.
RC STRAIN=MU50, and N315;
RX MEDLINE=21311952; PubMed=11418146;
RA Kuroda M., Ohta T., Uchiyama I., Baba T., Yuzawa H., Kobayashi I.,
RA Cui L., Oguchi A., Aoki K.-I., Nagai Y., Lian J.-Q., Ito T.,
RA Kanamori M., Matsumaru H., Maruyama A., Murakami H., Hosoyama A.,
RA Mizutani-Ui Y., Takahashi N.K., Sawano T., Inoue R.-I., Kaito C.,
RA Sekinizu K., Hirakawa H., Kuhara S., Goto S., Yabuzaki J.,
RA Kanehisa M., Yamashita A., Oshima K., Furuya K., Yoshino C., Shiba T.,
RA Hattori M., Ogasawara N., Hayashi H., Hiramatsu K.;
RT "Whole genome sequencing of methicillin-resistant Staphylococcus
RT aureus.";
RL Lancet 357:1225-1240(2001).
RN [2]
RC SEQUENCE FROM N.A.
RC STRAIN=MW2;
RX MEDLINE=22040717; PubMed=12044378;
RA Baba T., Takeuchi F., Kuroda M., Yuzawa H., Aoki K.-I., Oguchi A.,

```

archa  
CBI

```

AC Q8E4Q1;
DT 01-MAR-2003 (TrEMBLrel. 23, Created)
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DE 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE Hypothetical protein.
GN G8S1350.
OS Streptococcus agalactiae (serotype III).
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC Streptococcus.
OX NCBI_TaxID=216495;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=NEW316 / Serotype III;
RX MEDLINE=2242508; PubMed=12354221;
RA Glaser P., Rusniok C., Buchrieser C., Chevalier F., Frangeul L.,
RA Meadek T., Zouine M., Couve E., Lalloui L., Poyart C., Trieu-Cuot P.,
RA Kunst F.;
RT "Genome sequence of Streptococcus agalactiae, a pathogen causing
RT invasive neonatal disease.";
RL Mol. Microbiol. 45:1499-1513(2002).
DR EMBL; AL766850; CAD47009.1; -.
DR Sagaliet; gbs1350; -.
KW Hypothetical protein; Complete proteome.
SQ SEQUENCE 99 AA; 11302 MW; B83CB8FC1AB03C53 CRC64;

Query Match 37.4%; Score 34; DB 16; Length 99;
Best Local Similarity 46.7%; Pred. No. 4.6e+02;
Matches 7; Conservative 4; Mismatches 4; Indels 0; Gaps 0;

Qy 5 YDALNVLMMNIISK 19
Db 78 YIALNLLILGLVK 92

RESULT 14
Q9HQ82 PRELIMINARY; PRT; 66 AA.
AC Q9HQ82;
DT 01-MAR-2001 (TrEMBLrel. 16, Created)
DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT 01-MAR-2001 (TrEMBLrel. 16, Last annotation update)
DE Vng1283h.
GN Vng1283h.
OS Halobacterium sp. (strain NRC-1).
OC Archaea; Euryarchaeota; Halobacteria; Halobacteriales;
OC Halobacteriaceae; Halobacterium.
OX NCBI_TaxID=64091;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=20504483; PubMed=11016950;
RA Ng W.V., Kennedy S.P., Mahairas G.G., Berquist B., Pan M.,
RA Shukla H.D., Lasky S.R., Baliga N.S., Thorsson V., Shrogha J.,
RA Swartzell S., Weir D., Hall J., Dahl T.A., Welti R., Goo Y.A.,
RA Leithauser B., Keller K., Cruz R., Danison M.J., Hough D.W.,
RA Maddocks D.G., Jablonki P.E., Krebs M.P., Angevine C.M., Dale H.,
RA Isenbarger T.A., Peck R.F., Pohlischroder M., Spudich J.L., Jung K.-H.,
RA Alam M., Freitas T., Hou S., Daniels C.J., Dennis P.P., Omer A.D.,
RA Ehardt H., Lowe T.M., Liang P., Riley M., Hood L., Dassarma S.;
RT "Genome sequence of Halobacterium species NRC-1.";
RL Proc. Natl. Acad. Sci. U.S.A. 97:12176-12181(2000).
DR EMBL; AE005052; AAG19634.1; -.
KW Complete proteome.
SQ SEQUENCE 66 AA; 6859 MW; B5CD27577F80E8A8 CRC64;

Query Match 36.3%; Score 33; DB 17; Length 66;
Best Local Similarity 50.0%; Pred. No. 4.5e+02;
Matches 8; Conservative 4; Mismatches 4; Indels 0; Gaps 0;

Qy 3 RYVDALNVLMMNIIS 18
Db 20 RYVDALNVLGPVIVA 35

AC Q979G1 PRELIMINARY; PRT; 66 AA.
AC Q979G1;
DT 01-OCT-2001 (TrEMBLrel. 18, Created)
DT 01-OCT-2001 (TrEMBLrel. 18, Last sequence update)
DT 01-MAR-2002 (TrEMBLrel. 20, Last annotation update)
DE Hypothetical protein TV1200.
GN TV1200 OR TVG1229743.
OS Thermoplasma volcanium.
OC Archaea; Euryarchaeota; Thermoplasmata; Thermoplasmatales;
OC Thermoplasmataceae; Thermoplasma.
OX NCBI_TaxID=50339;
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN=GSS1 / DSM 4299 / JCM 9571;
RX MEDLINE=20570466; PubMed=11121031;
RA Kawashima T., Anano N., Koike H., Makino S.-I., Higuchi S.,
RA Kawashima T., Yamamoto Y., Watanabe K., Yamazaki M., Kanehori K., Kawamoto T.,
RA Nunoshiba T., Yamamoto Y., Aramaki H., Makino K., Suzuki M.;
RT "Archaeal adaptation to higher temperatures revealed by genomic
RT sequence of Thermoplasma volcanium.";
RL Proc. Natl. Acad. Sci. U.S.A. 97:14257-14262(2000).
DR EMBL; AP000995; BAB60342.1; -.
KW Hypothetical protein; Complete proteome.
SQ SEQUENCE 66 AA; 7786 MW; 7237A16A61E309FB CRC64;

Query Match 36.3%; Score 33; DB 17; Length 66;
Best Local Similarity 50.0%; Pred. No. 4.5e+02;
Matches 6; Conservative 4; Mismatches 2; Indels 0; Gaps 0;

Qy 8 LNLVLMNMNIISK 19
Db 50 LNIIVDKNIISR 61

RESULT 16
Q91038 PRELIMINARY; PRT; 79 AA.
AC Q91038;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-MAR-2002 (TrEMBLrel. 20, Last annotation update)
DE Ornithine decarboxylase (Fragment).
OS Gadus morhua (Atlantic cod).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
OC Acanthomorpha; Paracanthopterygii; Gadiformes; Gadidae; Gadus.
OX NCBI_TaxID=8049;
RN [1]
RP SEQUENCE FROM N.A.
RX Ong T.L., McNamara P.T., Armstrong R.F., Buckley L.J.;
RL Submitted (FEB-1996) to the EMBL/GenBank/DBJ databases.
DR EMBL; U49122; AAA91234.1; -.
DR HSSP; P11926; ID7K.
DR InterPro; IPR000183; Decarboxylase2.
DR Pfam; PF02784; Orn_Arg_dec_N; 1.
DR Pfam; PF00278; Orn_DAP_Arg_dec; 1.
FT NON_TER 1
SQ SEQUENCE 79 AA; 8919 MW; C3PEFF54CF0BCF2D CRC64;

Query Match 36.3%; Score 33; DB 13; Length 79;
Best Local Similarity 47.1%; Pred. No. 5.4e+02;
Matches 8; Conservative 3; Mismatches 6; Indels 0; Gaps 0;

Qy 3 RYVDALNVLMMNIISK 19
Db 32 RYTVASAYTLAVNIIAK 48

RESULT 17
Q8EIV8 PRELIMINARY; PRT; 83 AA.
ID Q8EIV8

```

Q8EIV8;  
 CC 01-MAR-2003 (T-EMBLrel. 23, Created)  
 CC 01-MAR-2003 (T-EMBLrel. 23, Last sequence update)  
 DT 01-MAR-2003 (T-EMBLrel. 23, Last annotation update)  
 DT 01-MAR-2003 (T-EMBLrel. 23, Last annotation update)  
 DE Conserved hypothetical protein.  
 GN S00721.  
 OS Shewanella oneidensis.  
 CC Bacteria; Proteobacteria; Gammaproteobacteria; Alteromonadales;  
 CC Alteromonadaceae; Shewanella.  
 CC NCBI\_TaxID=70863;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=WB-1;  
 RX MEDLINE=22297686; PubMed=12368813;  
 RA Heidelberg J.F., Paulsen I.T., Nelson K.E., Gaidos E.J., Nelson W.C.,  
 RA Read T.D., Eisen J.A., Sehadri R., Ward N., Methe B., Clayton R.A.,  
 RA Meyer T., Teapin A., Scott J., Beanan M., Brinkak L., Daugherty S.,  
 RA DeBoy R.T., Dodson R.J., Durkin A.S., Haft D.H., Kolonay J.F.,  
 RA Madupu R., Peterson J.D., Umavam L.A., White O., Wolf A.M.,  
 RA Vamathevan J., Weidman J., Impraim M., Lee K., Berry K., Lee C.,  
 RA Mueller J., Khouri H., Gill J., Utterback T.R., McDonald L.A.,  
 RA Feldblyum T.V., Smith H.O., Venter J.C., Nealeson K.H., Fraser C.M.;  
 RT "Genome sequence of the dissimilatory metal ion-reducing bacterium  
 RL Nat. Biotechnol. 20:1118-1123(2002).  
 DR EMBL; AE015517; AAN53799.1; --  
 DR TIGR; S00721; --  
 KW Hypothetical protein; Complete proteome.  
 SQ SEQUENCE 83 AA; 9075 MW; AC5D08F38ACB345C CRC64;  
 Query Match 36.3%; Score 33; DB 16; Length 83;  
 Best Local Similarity 45.08; Pred. NO. 5.7e+02;  
 Matches 9; Conservative 4; Mismatches 5; Indels 2; Gaps 1;  
 QY 2 RRVDAL--NVLMAMNIISK 19  
 Db 19 QELFQALTDNPLMANGIIGQ 38  
 RESULT 18  
 ID 063454 PRELIMINARY; PRT; 92 AA.  
 AC 063454;  
 DT 01-AUG-1998 (T-EMBLrel. 07, Created)  
 DT 01-AUG-1998 (T-EMBLrel. 07, Last sequence update)  
 DT 01-OCT-2002 (T-EMBLrel. 22, Last annotation update)  
 DE Cytochrome oxidase subunit I (EC 1.9.3.1) (Cytochrome c oxidase  
 DE polypeptide I) (Fragment).  
 OS Cupha erymanthis.  
 CC Mitochondrion.  
 CC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;  
 CC Neoptera; Endopterygota; Lepidoptera; Glossata; Ditrysia;  
 CC Papilionoidea; Nymphalidae; Heliconiinae; Cupha.  
 CC NCBI\_TaxID=64452;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=98119519; PubMed=9459431;  
 RA Brower A.V., DeSalle R.;  
 RT "Patterns of mitochondrial versus nuclear DNA sequence divergence  
 RT among nymphalid butterflies: the utility of wingless as a source of  
 RT characters for phylogenetic inference";  
 RL Insect Mol. Biol. 7:73-82(1998).  
 CC -!- FUNCTION: CYTOCHROME C OXIDASE IS THE COMPONENT OF THE RESPIRATORY  
 CC CHAIN THAT CATALYZES THE REDUCTION OF OXYGEN TO WATER. SUBUNIT 1-  
 CC 3 FORM THE FUNCTIONAL CORE OF THE ENZYME COMPLEX. CO I IS THE  
 CC CATALYTIC SUBUNIT OF THE ENZYME. ELECTRONS ORIGINATING IN  
 CC CYTOCHROME C ARE TRANSFERRED VIA THE COPPER A CENTER OF SUBUNIT 2  
 CC AND HEME A OF SUBUNIT 1 TO THE BIMETALLIC CENTER FORMED BY HEME A3  
 CC AND COPPER B (BY SIMILARITY).  
 CC -!- CATALYTIC ACTIVITY: 4 FERROCYTOCHROME C + O(2) = 4 FERROCYTOCHROME  
 CC C + 2 H2O.  
 CC -!- PATHWAY: RESPIRATORY CHAIN; TERMINAL STEP.  
 CC -!- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN. MITOCHONDRIAL

CC INNER MEMBRANE (BY SIMILARITY).  
 CC -!- SIMILARITY: BELONGS TO THE HEME-COPPER RESPIRATORY OXIDASE FAMILY.  
 DR EMBL; AF014151; AAC05893.1; --  
 DR InterPro; IPR000883; COX1.  
 DR Pfam; PF00115; COX1; 1.  
 KW Copper; Electron transport; Heme; Inner membrane; Membrane;  
 KW Oxidoreductase; Respiratory chain; Transmembrane; Transport;  
 KW Mitochondrion.  
 FT NON\_TER 1 1  
 SQ SEQUENCE 92 AA; 10826 MW; B7B6BEDFFDCE3697 CRC64;  
 Query Match 36.3%; Score 33; DB 8; Length 92;  
 Best Local Similarity 47.1%; Pred. NO. 6.3e+02;  
 Matches 8; Conservative 3; Mismatches 6; Indels 0; Gaps 0;  
 QY 2 RRVDALNVLMAMNIIS 18  
 Db 16 RRYSDYDPMFMSWNIIS 32  
 RESULT 19  
 ID 08XYL3 PRELIMINARY; PRT; 94 AA.  
 AC 08XYL3;  
 DT 01-MAR-2002 (T-EMBLrel. 20, Created)  
 DT 01-MAR-2002 (T-EMBLrel. 20, Last sequence update)  
 DT 01-JUN-2002 (T-EMBLrel. 21, Last annotation update)  
 DE Hypothetical protein RSC1745.  
 GN RSC1745 OR RS02933.  
 OS Ralstonia solanacearum (Pseudomonas solanacearum).  
 CC Bacteria; Proteobacteria; Betaproteobacteria; Burkholderiales;  
 CC Ralstoniaceae; Ralstonia.  
 CC NCBI\_TaxID=305;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=GM11000;  
 RX MEDLINE=21681879; PubMed=11823852;  
 RA Salanoubat M., Genin S., Artiguenave F., Gouzy J., Mangenot S.,  
 RA Arlat M., Billault A., Brottier P., Camus J.C., Catolico L.,  
 RA Chandler M., Choiane N., Claudel-Renard C., Cunnac S., Demange N.,  
 RA Gaspin C., Lavie M., Moisan A., Robert C., Saurin W., Schiex T.,  
 RA Signier P., Thebault P., Whalen M., Wincker P., Levy M.,  
 RA Weissenbach J., Boucher C.A.;  
 RT "Genome sequence of the plant pathogen Ralstonia solanacearum";  
 RL Nature 415:497-502(2002).  
 DR EMBL; AL646066; CAD15447.1; --  
 DR InterPro; IPR002634; BOLA.  
 DR Pfam; PF01722; BOLA; 1.  
 KW Hypothetical protein; Complete proteome.  
 SQ SEQUENCE 94 AA; 10195 MW; C696918D04E01A70 CRC64;  
 Query Match 36.3%; Score 33; DB 16; Length 94;  
 Best Local Similarity 70.0%; Pred. NO. 6.4e+02;  
 Matches 7; Conservative 0; Mismatches 3; Indels 0; Gaps 0;  
 QY 2 RRVDALNVLL 11  
 Db 66 RMVYDALRTL 75  
 RESULT 20  
 ID 096DE8 PRELIMINARY; PRT; 95 AA.  
 AC 096DE8;  
 DT 01-DEC-2001 (T-EMBLrel. 19, Created)  
 DT 01-DEC-2001 (T-EMBLrel. 19, Last sequence update)  
 DT 01-OCT-2002 (T-EMBLrel. 22, Last annotation update)  
 DE Similar to 26S proteasome-associated pad1 homolog.  
 OS Homo sapiens (Human).  
 CC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 CC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 CC NCBI\_TaxID=9606;  
 RN [1]

RP SEQUENCE FROM N.A.  
 RC TISSUE=Lung;  
 RA Strausberg R.;  
 RL Submitted (JUN-2001) to the EMBL/GenBank/DBJ databases.  
 DR EMBL: BC009524; AA09524.1; -  
 DR InterPro: IPR003639; Mov34\_1.  
 DR ProDom: PD363422; Mov34\_1; 1.  
 KW Proteasome.  
 SQ SEQUENCE 95 AA; 11009 MW; D70270AD00529A18 CRC64;

Query Match 36.3%; Score 33; DB 4; Length 95;  
 Best Local Similarity 25.3%; Pred. No. 6.5e+02;  
 Matches 6; Conservative 5; Mismatches 5; Indels 0; Gaps 0;

Qy 1 RRVYDALNVLMMNII 17  
 :|: : :|:|:|:|:|:  
 Db 66 KRHLEHVVDVLMISNIV 82

## RESULT 21

Q8HA80 PRELIMINARY; PRT; 89 AA.  
 AC Q8HA80;  
 DT 01-MAR-2003 (TrEMBLrel. 23, Created)  
 DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)  
 DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)  
 DE Hypothetical protein.  
 GN ORF2C.  
 OS Bacteriophage T5.  
 OC Viruses; dsDNA viruses, no RNA stage; Caudovirales; Siphoviridae;  
 OC T5-like viruses.  
 OC NCBI\_TaxID=10726;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Mikoulinskaia G.V., Zimin A.A., Feofanov S.A.;  
 RL Submitted (AUG-2002) to the EMBL/GenBank/DBJ databases.  
 DR EMBL: AY140897; AANL7754.1; -  
 KW Hypothetical protein.  
 SQ SEQUENCE 89 AA; 10284 MW; 59C7EF0AC84D9B3C CRC64;

Query Match 35.7%; Score 32.5; DB 9; Length 89;  
 Best Local Similarity 43.8%; Pred. No. 7.4e+02;  
 Matches 7; Conservative 5; Mismatches 3; Indels 1; Gaps 1;

Qy 2 RRVYDALNV-LMAMNI 16  
 :|:|:|:|:|:|:|:  
 Db 41 KMYDLSLNKMAFNL 56

## RESULT 22

O44050 PRELIMINARY; PRT; 24 AA.  
 AC O44050;  
 DT 01-JUN-1998 (TrEMBLrel. 06, Created)  
 DT 01-JUN-1998 (TrEMBLrel. 06, Last sequence update)  
 DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)  
 DE Sialidase homolog (Fragment).  
 OS Trypanosoma rangeli.  
 OC Eukaryota; Euklenozoa; Kinetoplastida; Trypanosomatidae; Trypanosoma.  
 OC NCBI\_TaxID=5698;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=LDS;  
 RX MEDLINE=98115452; PubMed=9455917;  
 RA Buechiazzo A., Campetella O., Frasch A.C.C.;  
 RT "Trypanosoma rangeli sialidase: cloning, expression and similarity to  
 T. cruzi trans-sialidase."  
 RL Glycobiology 7:1167-1173(1997).  
 DR EMBL: U83240; AAC38897.1; -  
 FT NON\_TER 1  
 FT NON\_TER 24  
 SQ SEQUENCE 24 AA; 2695 MW; 009FE2472BBD3F35 CRC64;

Query Match 35.2%; Score 32; DB 5; Length 24;  
 Best Local Similarity 46.2%; Pred. No. 2.5e+02;  
 Matches 6; Conservative 4; Mismatches 3; Indels 0; Gaps 0;

Qy 1 RRVYDALNVLMA 13  
 | | | | :|:|:|:  
 Db 12 RERVDSFVILLS 24

## RESULT 23

O8X3F6 PRELIMINARY; PRT; 53 AA.  
 AC O8X3F6;  
 DT 01-MAR-2002 (TrEMBLrel. 20, Created)  
 DT 01-MAR-2002 (TrEMBLrel. 20, Last sequence update)  
 DT 01-MAR-2002 (TrEMBLrel. 20, Last annotation update)  
 DE Hypothetical protein EC83484.  
 GN EC83484.  
 OS Escherichia coli O157:H7.  
 OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;  
 OC Enterobacteriaceae; Escherichia.  
 OC NCBI\_TaxID=83334;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=O157:H7 / RIMD 0509952;  
 RX MEDLINE=21156231; PubMed=11258796;  
 RA Hayashi T., Makino K., Ohnishi M., Kurokawa K., Ishii K., Yokoyama K.,  
 Han C.-G., Ohtsubo E., Nakayama K., Murata T., Tanaka M., Tobe T.,  
 Iida T., Takami H., Honda T., Sasakawa C., Ogasawara N., Yasunaga T.,  
 Kuhara S., Shiba T., Hattori M., Shingawa H.;  
 RA "Complete genome sequence of enterohemorrhagic Escherichia coli  
 O157:H7 and genomic comparison with a laboratory strain K-12."  
 RL DNA Res. 8:11-22(2001).  
 DR EMBL: AP002562; BAB36907.1; -  
 KW Hypothetical protein.  
 SQ SEQUENCE 53 AA; 6018 MW; 653FBE5407E4B98F CRC64;

Query Match 35.2%; Score 32; DB 16; Length 53;  
 Best Local Similarity 33.3%; Pred. No. 5.4e+02;  
 Matches 6; Conservative 5; Mismatches 7; Indels 0; Gaps 0;

Qy 2 RRVYDALNVLMMNIIISK 19  
 | :|:|:|:|:|:|:  
 Db 36 RTTQTSONIVAAMNLYNK 53

## RESULT 24

P92499 PRELIMINARY; PRT; 55 AA.  
 AC P92499;  
 DT 01-MAY-1997 (TrEMBLrel. 03, Created)  
 DT 01-MAY-1997 (TrEMBLrel. 03, Last sequence update)  
 DT 01-OCT-2002 (TrEMBLrel. 22, Last annotation update)  
 DE Cytochrome oxidase subunit III (EC 1.9.3.1) (Cytochrome c oxidase  
 polypeptide III) (Fragment).  
 GN COIII.  
 OS Apis mellifera (Honeybee).  
 OC Mitochondrion.  
 OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;  
 OC Neoptera; Endopterygota; Hymenoptera; Apocrita; Aculeata; Apoidea;  
 OC Apidae; Apis.  
 OC NCBI\_TaxID=7460;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Koulianos S., Crozier R.H.;  
 RT "Mitochondrial sequence characterisation of Australian commercial and  
 feral honeybee strains (Hymenoptera: Apidae: Apis mellifera Linnaeus),  
 RT in the context of the species worldwide."  
 RL J. Aust. Entomol. Soc. 0:0-0(1997).  
 CC -1- FUNCTION: SUBUNITS I, II AND III FORM THE FUNCTIONAL CORE OF THE  
 CC ENZYME COMPLEX (BY SIMILARITY).  
 CC -1- CATALYTIC ACTIVITY: 4 FERROCYTOCHROME C + O(2) = 4 FERRICYTOCHROME  
 CC C + 2 H(2)O.

GN	MTH186.
OS	Methanobacterium thermoautotrophicum.
OC	Archaea; Euryarchaeota; Methanobacteria; Methanobacteriales;
OC	Methanobacteriaceae; Methanothermobacter.
NCBI_TaxID=187420;	
OX	[1]
RN	SEQUENCE FROM N.A.
RP	STRAIN=Delta H;
RC	MEDLINE=98037514; PubMed=9371463;
RX	Smith D.R., Doucette-Stamm L.A., DeLoughery C., Lee H.-M., Dubois J.,
RA	Aldredge T., Bashirzadeh R., Blakely D., Cook R., Gilbert K.,
RA	Harrison D., Hoang L., Keagle P., Lum W., Pothier B., Qiu D.,
RA	Spadafora R., Vicare A.R., Wang Y., Wierzbowski J., Gibson R.,
RA	Jiwani N., Caruso A., Bush D., Safer H., Patwell D., Prabhakar S.,
RA	McDougall S., Shimer G., Goyal A., Pietrowski S., Church G.M.,
RA	Daniels C.J., Mao J.-I., Rice P., Noelling J., Reeve J.N.;
RT	"Complete genome sequence of Methanobacterium thermoautotrophicum
RT	deltah; functional analysis and comparative genomics.";
RL	J. Bacteriol. 179:7135-7155(1997).
KW	EMBL; AF000806; AAB84692.1; -
DW	Hypothetical protein; Complete proteome.
KX	HYPOTHETICAL PROTEIN; Complete proteome.
SQ	SEQUENCE 60 AA; 6960 MW; A3FACA6838251B4D CRC64;
Query Match	35.2%; Score 32; DB 17; Length 60;
Best Local Similarity	46.2%; Pred. No. 6.1e+02;
Matches	6; Conservative 4; Mismatches 3; Indels 0; Gaps 0;
QY	7 ALNVLMAMNIISK 19
DB	: : : : : :
	20 AINISIFLNICK 32
RESULT 27	
Q9Z682	PRELIMINARY; PRT; 61 AA.
ID	Q9Z682
AC	Q9Z682;
DT	01-MAY-1999 (TrEMBLrel. 10, Created)
DT	01-MAY-1999 (TrEMBLrel. 10, Last sequence update)
DT	01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE	Exonuclease I [Fragment].
OC	Bradyrhizobium japonicum.
OC	Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;
OC	Bradyrhizobiaceae; Bradyrhizobium.
OC	NCBI_TaxID=375;
RN	[1]
RC	SEQUENCE FROM N.A.
RC	STRAIN=USDALL0;
RT	Mayer R.M., Mathis J.N., McMillin D.E.;
RT	"Analysis of a DNA fragment present in Bradyrhizobium japonicum 110
RT	Fix+ derivatives but missing in a Fix- derivative.";
RL	Submitted (OCT-1998) to the EMBL/GenBank/DBJ databases.
DR	EMBL; AF101073; AAD17890.1; -
FT	NON TER 61
SQ	SEQUENCE 61 AA; 6729 MW; FF29A9A1891C986A CRC64;
Query Match	35.2%; Score 32; DB 2; Length 61;
Best Local Similarity	46.2%; Pred. No. 6.2e+02;
Matches	6; Conservative 5; Mismatches 2; Indels 0; Gaps 0;
QY	6 DALNVLMAMNIIS 18
DB	: : : : : :
	25 DALSLVMAADAVS 37
RESULT 28	
Q8UKS1	PRELIMINARY; PRT; 64 AA.
ID	Q8UKS1
AC	Q8UKS1;
DT	01-OCT-2002 (TrEMBLrel. 22, Created)
DT	01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
DT	01-OCT-2002 (TrEMBLrel. 22, Last annotation update)
DE	Arginine-rich protein a209R.
GN	ORF40.



```

OS Heliothis zea virus 1.
OC Viruses; unclassified viruses.
OX NCBI_TaxID=29250;
RN [1]
RX MEDLINE=22174892; PubMed=12186886;
RA Cheng C.H., Liu S.M., Chow T.Y., Hsiao Y.Y., Wang D.P., Huang J.J.,
RA Chen H.H.;
RT "Analysis of the Complete Genome Sequence of the H2-1 Virus Suggests
RT that it is Related to Members of the Baculoviridae.";
RN J. Virol. 76:9024-9034(2002).
RN [2]
RP SEQUENCE FROM N.A.
RA Cheng C.H., Liu H.M., Hsiao Y.Y., Chow T.Y., Chen H.H.;
RL Submitted (JUN-2001) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RA Cheng C.H., Hsiao Y.Y., Liu S.M., Chow T.Y., Chen H.H.;
RL Submitted (JUN-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF451898; AAN04335.1; -
SQ SEQUENCE 64 AA; 7698 MW; 775234CECB423428 CRC64;
Query Match 35.2%; Score 32; DB 12; Length 64;
Best Local Similarity 31.6%; Pred. No. 6.5e+02;
Matches 6; Conservative 4; Mismatches 9; Indels 0; Gaps 0;
QY 1 RRVYDALNVLMMNIISK 19
Db 39 RKNVLQKNVLQKNVLQR 57
RESULT 29
Q99RD1 PRELIMINARY; PRT; 66 AA.
AC Q99RD1;
DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DE Hypothetical protein SAV2504 (Hypothetical protein MW2422).
GN SAV2504 OR SA2292 OR MW2422.
OS Staphylococcus aureus (strain N315), and
OS Staphylococcus aureus (strain N315), and
OS Staphylococcus aureus (strain N315), and
OC Bacteria; Firmicutes; Bacillales; Staphylococcus.
OX NCBI_TaxID=158878, 158879, 196620;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=M50, and N315;
RX MEDLINE=21311952; PubMed=11418146;
RA Kuroda M., Ohta T., Uchiyama I., Baba T., Yuzawa H., Kobayashi I.,
RA Cui L., Oguchi A., Aoki K.-I., Nagai Y., Lian J.-Q., Ito T.,
RA Kanamori M., Matsumaru H., Maruyama A., Murakami H., Hosoyama A.,
RA Mizunari-Ui Y., Takahashi N.K., Sawano T., Inoue R.-I., Kaito C.,
RA Sekimizu K., Hirakawa H., Kuhara S., Goto S., Yabuzaki J.,
RA Kanehisa M., Yamashita A., Ohnishi K., Furuya K., Yoshino C., Shiba T.,
RA Hattori M., Ogasawara N., Hayashi H., Hiramatsu K.;
RT "Whole genome sequencing of methicillin-resistant Staphylococcus
RT aureus.";
RL Lancet 357:1225-1240(2001).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=MW2;
RX MEDLINE=22040717; PubMed=12044378;
RA Baba T., Takeuchi F., Kuroda M., Yuzawa H., Aoki K.-I., Oguchi A.,
RA Nagai Y., Iwama N., Asano K., Naimi T., Kuroda H., Cui L.,
RA Yamamoto K., Hiramatsu K.;
RT "Genome and virulence determinants of high virulence community-
RT acquired MRSA.";
RL Lancet 359:1819-1827(2002).
DR EMBL; AP003365; BAB58666.1; -
DR EMBL; AP003337; BAB43595.1; -
DR EMBL; AP004630; BAB96287.1; -
KW Hypothetical protein; Complete proteome.

```

---

```

SQ SEQUENCE 66 AA; 8016 MW; 8B6BB4B3627F791C CRC64;
Query Match 35.2%; Score 32; DB 16; Length 66;
Best Local Similarity 66.7%; Pred. No. 6.8e+02;
Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;
QY 3 RYVDALNVL 11
Db 47 RYVDLNTI 55
RESULT 30
Q9GMS9 PRELIMINARY; PRT; 71 AA.
AC Q9GMS9;
DT 01-MAR-2001 (TrEMBLrel. 16, Created)
DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT 01-MAR-2001 (TrEMBLrel. 16, Last annotation update)
DE Hypothetical 8.3 kDa protein.
OS Macaca fascicularis (Crab eating macaque) (Cynomolgus monkey).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Euthera; Primates; Catarrhini; Cercopitheciidae;
OC Cercopitheciinae; Macaca.
OX NCBI_TaxID=9541;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Cerebellum;
RA Osada N., Hida M., Kusuda J., Tanuma R., Iseki K., Hirai M., Terao K.,
RA Suzuki Y., Sugano S., Hashimoto K.;
RT "Isolation of full-length cDNA clones from macaque brain cDNA
RT libraries.";
RL Submitted (AUG-2000) to the EMBL/GenBank/DBJ databases.
KW EMBL; AB047839; BAB12265.1; -
SQ SEQUENCE 71 AA; 8296 MW; 1A50C4AE703A69C7 CRC64;
Query Match 35.2%; Score 32; DB 6; Length 71;
Best Local Similarity 33.3%; Pred. No. 7.3e+02;
Matches 5; Conservative 5; Mismatches 5; Indels 0; Gaps 0;
QY 2 RRVYDALNVLMMNI 16
Db 9 KKLVDNINVKMLSQV 23
RESULT 31
Q94EX1 PRELIMINARY; PRT; 72 AA.
AC Q94EX1;
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE A74924411.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eurosids II; Brassicales; Brassicaceae; Arabidopsis.
OX NCBI_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.
RA Kim C.J., Chen H., Cheuk R., Koesema E., Meyers M.C., Shinn P.,
RA Tracy S.B., Banh J., Bowser L., Carninci P., Chung M.K.,
RA Goldsmith A.D., Hayashizaki Y., Ishida J., Jones T., Kamiya A.,
RA Karlin-Neumann G., Kawai J., Lam B., Lee J.M., Lin J., Liu S.X.,
RA Miranda M., Narusaka M., Nguyen M., Palm C.J., Pham P.K., Quach H.L.,
RA Sakano H., Sakurai T., Satou M., Seki M., Southwick A., Tang C.C.,
RA Toriumi M., Yamada K., Yu G., Shinozaki K., Davis R.W., Theologis A.,
RA Ecker J.R.;
RT "Arabidopsis cDNA clones.";
RL Submitted (JUN-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF389288; AAK63860.1; -
SQ SEQUENCE 72 AA; 8265 MW; C0CA73AC066EC1C2 CRC64;

```

Query Match		35.2%;	Score 32;	DB 10;	Length 72;	
Best Local Similarity		60.0%;	Pred. No. 7.4e+02;			
Matches		6;	Conservative	3;	Mismatches	1; Indels 0; Gaps 0;
QY	10 VLMAMNISK 19	: : : :				
Db	33 VMVAKNIVSK 42	: : : :				
RESULT 32						
ID	Q34371	PRELIMINARY;	PRT;	77 AA.		
AC	Q34371					
DT	01-NOV-1996 (TrEMBLrel. 01, Created)					
DT	01-NOV-1996 (TrEMBLrel. 01, Last sequence update)					
DT	01-OCT-2002 (TrEMBLrel. 22, Last annotation update)					
DE	Cytochrome oxidase I (EC 1.9.3.1) (Cytochrome c oxidase polypeptide I) (Fragment).					
DE	Fragment.					
OS	Dryadula phaetusa.					
OG	Mitochondrion.					
OC	Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;					
OC	Neoptera; Endopterygota; Lepidoptera; Glossata; Ditrysia;					
OC	Papilionoidea; Nymphalidae; Heliconiinae; Dryadula.					
OX	NCBI_TaxID=34742;					
RN	[1]					
RP	SEQUENCE FROM N.A.					
RC	STRAIN=P-11-7; TISSUE=Head, and Thorax;					
RX	MEDLINE=94356263; PubMed=8075834;					
RA	Brower A.V.Z.;					
RA	"Phylogeny of Heliconius butterflies inferred from mitochondrial DNA sequences (Lepidoptera: Nymphalidae).";					
RL	Mol. Phylogenet. Evol. 3:159-174(1994).					
CC	-1- FUNCTION: CYTOCHROME C OXIDASE IS THE COMPONENT OF THE RESPIRATORY CHAIN THAT CATALYZES THE REDUCTION OF OXYGEN TO WATER. SUBUNITS 1-3 FORM THE FUNCTIONAL CORE OF THE ENZYME COMPLEX. CO I IS THE CATALYTIC SUBUNIT OF THE ENZYME. ELECTRONS ORIGINATING IN CYTOCHROME C ARE TRANSFERRED VIA THE COPPER A CENTER OF SUBUNIT 2 AND HEME A OF SUBUNIT 1 TO THE BIMETALLIC CENTER FORMED BY HEME A3 AND COPPER B (BY SIMILARITY).					
CC	-1- CATALYTIC ACTIVITY: 4 FERROCYTOCHROME C + O(2) = 4 FERRICYTOCHROME C + 2 H(2)O.					
CC	-1- PATHWAY: RESPIRATORY CHAIN; TERMINAL STEP.					
CC	-1- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN. MITOCHONDRIAL INNER MEMBRANE (BY SIMILARITY).					
CC	-1- SIMILARITY: BELONGS TO THE HEME-COPPER RESPIRATORY OXIDASE FAMILY.					
DR	EMBL; U08546; AAA20722.1; -.					
DR	InterPro; IPR00883; COX1.					
DR	Pfam; PF00115; COX1; 1.					
KW	Copper; Electron transport; Heme; Inner membrane; Membrane; Oxidoreductase; Respiratory chain; Transmembrane; Transport; Mitochondrion.					
FT	NON_TER					
SQ	SEQUENCE 77 AA; 9133 MW; 7CB02EB82BBCC12D CRC64;					
Query Match						
Best Local Similarity		35.2%;	Score 32;	DB 8;	Length 77;	
Matches		8;	Conservative	2;	Mismatches	7; Indels 0; Gaps 0;
QY	2 RRVYDALNVLMMNIIIS 18	: : : :				
Db	1 RRYSDYDPDSFMWNVIS 17	: : : :				
RESULT 33						
ID	O57412	PRELIMINARY;	PRT;	78 AA.		
AC	O57412;					
DT	01-JUN-1998 (TrEMBLrel. 06, Created)					
DT	01-JUN-1998 (TrEMBLrel. 06, Last sequence update)					
DT	01-JUN-2002 (TrEMBLrel. 21, Last annotation update)					
DE	Guanine nucleotide dissociation stimulator (Fragment).					
GN	GNDS.					
OS	Tetraodon fluviatilis (Puffer fish).					
Query Match						
Best Local Similarity		35.2%;	Score 32;	DB 16;	Length 83;	
Matches		6;	Conservative	1;	Mismatches	3; Indels 0; Gaps 0;
QY	2 RRVYDALNVL 11	: : : :				
Db	14 REIYDTLNG 23	: : : :				
RESULT 35						
ID	Q54418	PRELIMINARY;	PRT;	89 AA.		
AC	Q54418;					
DT	01-NOV-1996 (TrEMBLrel. 01, Created)					
DT	01-NOV-1996 (TrEMBLrel. 01, Last sequence update)					
DT	01-NOV-1998 (TrEMBLrel. 08, Last annotation update)					
DE	Extracellular secretory protein.					
GN	NUCE.					
Query Match						
Best Local Similarity		60.0%;	Pred. No. 8.5e+02;			
Matches		6;	Conservative	1;	Mismatches	3; Indels 0; Gaps 0;
QY	3 RRYVDALNVLMMN 15	: : : :				
Db	25 RIPDNANVFYAMN 37	: : : :				
RESULT 34						
ID	Q9K761	PRELIMINARY;	PRT;	83 AA.		
AC	Q9K761;					
DT	01-OCT-2000 (TrEMBLrel. 15, Created)					
DT	01-OCT-2000 (TrEMBLrel. 15, Last sequence update)					
DT	01-MAR-2002 (TrEMBLrel. 20, Last annotation update)					
DE	Hypothetical protein BH3512.					
GN	BH3512.					
OS	Bacillus halodurans.					
OC	Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.					
OX	NCBI_TaxID=86665;					
RN	[1]					
RP	SEQUENCE FROM N.A.					
RC	STRAIN=C-125 / JCM 9153;					
RX	MEDLINE=20512582; PubMed=11058132;					
RA	Takami H., Nakasone K., Takaki Y., Maeno G., Sasaki R., Masui N., Fuji F., Hirama C., Nakamura Y., Ogasawara N., Kuhara S., Horikoshi K.;					
RA	"Complete genome sequence of the alkaliphilic bacterium Bacillus halodurans and genomic sequence comparison with Bacillus subtilis.";					
RL	Nucleic Acids Res. 28:4317-4331(2000).					
DR	EMBL; AP001519; BAB07231.1; -.					
KW	Hypothetical protein; Complete proteome.					
SQ	SEQUENCE 83 AA; 9454 MW; 471552D292B2A44F CRC64;					
Query Match						
Best Local Similarity		35.2%;	Score 32;	DB 13;	Length 78;	
Matches		7;	Conservative	1;	Mismatches	5; Indels 0; Gaps 0;
QY	3 RRYVDALNVLMMN 15	: : : :				
Db	25 RIPDNANVFYAMN 37	: : : :				
RESULT 36						
ID	Q9K761	PRELIMINARY;	PRT;	83 AA.		
AC	Q9K761;					
DT	01-OCT-2000 (TrEMBLrel. 15, Created)					
DT	01-OCT-2000 (TrEMBLrel. 15, Last sequence update)					
DT	01-MAR-2002 (TrEMBLrel. 20, Last annotation update)					
DE	Hypothetical protein BH3512.					
GN	BH3512.					
OS	Bacillus halodurans.					
OC	Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.					
OX	NCBI_TaxID=86665;					
RN	[1]					
RP	SEQUENCE FROM N.A.					
RC	STRAIN=C-125 / JCM 9153;					
RX	MEDLINE=20512582; PubMed=11058132;					
RA	Takami H., Nakasone K., Takaki Y., Maeno G., Sasaki R., Masui N., Fuji F., Hirama C., Nakamura Y., Ogasawara N., Kuhara S., Horikoshi K.;					
RA	"Complete genome sequence of the alkaliphilic bacterium Bacillus halodurans and genomic sequence comparison with Bacillus subtilis.";					
RL	Nucleic Acids Res. 28:4317-4331(2000).					
DR	EMBL; AP001519; BAB07231.1; -.					
KW	Hypothetical protein; Complete proteome.					
SQ	SEQUENCE 83 AA; 9454 MW; 471552D292B2A44F CRC64;					

```

OS Serratia marcescens.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Serratia.
OX NCBI_TaxID=615;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=SM6;
RX MEDLINE=96174474; PubMed=8594195;
RA Jin S., Chen Y., Christie G.E., Benedik M.J.;
RT "Regulation of the Serratia marcescens extracellular nuclease:
RT positive control by a homolog of P2 Ogr encoded by a cryptic
RT prophage.";
RL J. Mol. Biol. 256:264-278 (1996).
DR EMBL; U11698; AAA98439.1; -.
SQ SEQUENCE 89 AA; 10283 MW; DDAAD9AB45DFEAA CRC64;

Query Match 35.2%; Score 32; DB 2; Length 89;
Best Local Similarity 55.6%; Pred. No. 9.1e+02;
Matches 5; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 1 RRVYDALN 9
Db 80 RKKYDECN 88

RESULT 36
Q9HRH0
ID Q9HRH0 PRELIMINARY; PRT; 94 AA.
AC Q9HRH0
DT 01-MAR-2001 (TrEMBLrel. 16, Created)
DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT 01-MAR-2001 (TrEMBLrel. 16, Last annotation update)
DE Vng0703h.
GN VNG0703H.
OS Halobacterium sp. (strain NRC-1).
OC Archaea; Euryarchaeota; Halobacteria; Halobacteriales;
OC Halobacteriaceae; Halobacterium.
OX NCBI_TaxID=64091;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=20504483; PubMed=11016950;
RA Ng W.V., Kennedy S.P., Mahairas G.G., Berquist B., Pan M.,
RA Shukla H.D., Leaky S.R., Baliga N.S., Thorsson V., Sbrogna J.,
RA Swartzell S., Weir D., Hall J., Dahl T.A., Welti R., Goo Y.A.,
RA Leithausen D.G., Jellison K., Cruz R., Danson M.J., Hough D.W.,
RA Maddocks D.G., Jablonski P.E., Krebs M.P., Angevine C.M., Dale H.,
RA Isenbarger T.A., Peck R.F., Pohlechröder M., Spudich J.L., Jung K.-H.,
RA Alam M., Freitas T., Hou S., Daniels C.J., Dennis P.P., Omer A.D.,
RA Ehardt H., Lowe T.M., Liang P., Riley M., Hood L., DasSarma S.;
RT "Genome sequence of Halobacterium species NRC-1.";
RL Proc. Natl. Acad. Sci. U.S.A. 97:12176-12181(2000).
DR EMBL; AE005015; AAG19188.1; -.
KW Complete proteome.
SQ SEQUENCE 94 AA; 10601 MW; 4EA3976DDC6146B9 CRC64;

Query Match 35.2%; Score 32; DB 17; Length 94;
Best Local Similarity 36.8%; Pred. No. 9.6e+02;
Matches 7; Conservative 5; Mismatches 7; Indels 0; Gaps 0;

QY 1 RRVYDALNVMANNISK 19
Db 41 QRTVRDALDRLOADVVEK 59

RESULT 37
Q8V6V2
ID Q8V6V2 PRELIMINARY; PRT; 96 AA.
AC Q8V6V2
DT 01-MAR-2002 (TrEMBLrel. 20, Created)
DT 01-MAR-2002 (TrEMBLrel. 20, Last sequence update)
DT 01-MAR-2002 (TrEMBLrel. 20, Last annotation update)
DE Hypothetical 10.9 kDa protein.
OS Halovirus HF2.

```

```

OC Viruses; dsDNA viruses, no RNA stage; Caudovirales; Myoviridae.
OX NCBI_TaxID=33771;
RN [1]
RP SEQUENCE FROM N.A.
RA Tang S.-L., Fisher C., Ngui K., Nuttall S.D., Dyall-Smith M.L.;
RT "Sequence and transcription of halovirus HF2.";
RL Submitted (JAN-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF22060; AAL54928.1; -.
KW Hypothetical protein.
SQ SEQUENCE 96 AA; 10900 MW; 359249FBD43CF3CB CRC64;

Query Match 35.2%; Score 32; DB 12; Length 96;
Best Local Similarity 70.0%; Pred. No. 9.8e+02;
Matches 7; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 2 RRVYDALNVL 11
Db 87 RELEDALNVL 96

RESULT 38
O30267
ID O30267 PRELIMINARY; PRT; 96 AA.
AC O30267
DT 01-JAN-1998 (TrEMBLrel. 05, Created)
DT 01-JAN-1998 (TrEMBLrel. 05, Last sequence update)
DT 01-MAR-2002 (TrEMBLrel. 20, Last annotation update)
DE Hypothetical protein AF2404.
GN AF2404.
OS Archaeoglobus fulgidus.
OC Archaea; Euryarchaeota; Archaeoglobi; Archaeoglobales;
OC Archaeoglobaceae; Archaeoglobus.
OX NCBI_TaxID=2234;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=VC-16 / DSM 4304 / ATCC 49558;
RX MEDLINE=98049343; PubMed=9389475;
RA Klenk H.-P., Clayton R.A., Tomb J.-F., White O., Nelson K.E.,
RA Ketchum K.A., Dodson R.J., Winn M., Hickey E.K., Peterson J.D.,
RA Richardson D.L., Kerlavage A.R., Graham D.E., Kyrpides N.C.,
RA Fleischmann R.D., Quackenbush J., Lee N.H., Sutton G.G., Gill S.,
RA Kirkness E.F., Dougherty B.A., McKenney K., Adams M.D., Loftus B.,
RA Peterson S., Reich C.I., McNeil L.K., Badger J.H., Glodek A., Zhou L.,
RA Overbeek R., Gocayne J.D., Weidman J.F., McDonald L., Utterback T.,
RA Cotton M.D., Spriggs T., Artach P., Kaine B.P., Sykes S.M.,
RA Sadow P.M., D'Andrea K.P., Bowman C., Fujii C., Garland S.A.,
RA Mason T.M., Olsen G.J., Fraser C.M., Smith H.O., Woese C.R.,
RA Venter J.C.;
RT "The complete genome sequence of the hyperthermophilic, sulphate-
RT reducing archaeon Archaeoglobus fulgidus.";
RL Nature 390:364-370(1997).
CC -1- SIMILARITY: BELONGS TO THE ARSR FAMILY OF TRANSCRIPTIONAL
CC REGULATORS.
DR EMBL; AE001110; AAB91262.1; -.
DR TIGR; AF2404; -.
DR InterPro; IPR001845; HTH_ArsR.
DR Pfam; PF01022; HTH_5; 1.
KW DNA-binding; Hypothetical protein; Transcription regulation;
KW Complete proteome.
SQ SEQUENCE 96 AA; 10921 MW; E6AEB9A936E386A35 CRC64;

Query Match 35.2%; Score 32; DB 17; Length 96;
Best Local Similarity 50.0%; Pred. No. 9.8e+02;
Matches 6; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

QY 8 LNVLMANNISK 19
Db 60 LKALMQLNIVEK 71

RESULT 39
Q38507
ID Q38507 PRELIMINARY; PRT; 52 AA.

```

```
AC Q38507;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE Left end of bacteriophage phi-29 coding for 15 potential proteins
DE Among these are the terminal protein and the proteins encoded by the
DE genes 1, 2 (sus), 3, and (probably) 4.
OS Bacteriophage phi-29.
OC Viruses; dsDNA viruses, no RNA stage; Caudovirales; Podoviridae;
OC phi-29-like viruses.
OX NCBI_TaxID=10756;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=82262795; PubMed=6809534;
RA Yoshikawa H., Ito J.;
RT "Nucleotide sequence of the major early region of bacteriophage phi-29."
RL Gene 17:323-335(1982).
DR EMBL: V01155; CAA24488.1; B7E08BBB9ED0690 CRC64;
SQ SEQUENCE 52 AA; 5979 MW; B7E08BBB9ED0690 CRC64;

Query Match 34.6%; Score 31.5; DB 9; Length 52;
Best Local Similarity 43.8%; Pred. No. 6.5e+02;
Matches 7; Conservative 5; Mismatches 3; Indels 1; Gaps 1;

QY 2 RRVYDALNVLMMNII 17
DB 13 RDVFTSLNVI-RMNIV 27

RESULT 40
Q9KAUO PRELIMINARY; PRT; 35 AA.
AC Q9KAUO;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-MAR-2002 (TrEMBLrel. 20, Last annotation update)
DE Hypothetical protein BH2196.
GN BH2196.
OS Bacillus halodurans.
OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.
OX NCBI_TaxID=86665;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C-125 / JCM 9153;
RX MEDLINE=20512582; PubMed=11058132;
RA Takami H., Nakasone K., Takaki Y., Maeno G., Sasaki R., Masui N.,
RA Fuji F., Hirama C., Nakamura Y., Ogasawara N., Kuhara S.,
RA Horikoshi K.;
RT "Complete genome sequence of the alkaliphilic bacterium Bacillus
RT halodurans and genomic sequence comparison with Bacillus subtilis."
RL Nucleic Acids Res. 28:4317-4331(2000).
DR EMBL: AP001514; BAB05915.1; -
KW Hypothetical protein; Complete proteome.
SQ SEQUENCE 35 AA; 3960 MW; 9F851AF909DB2388 CRC64;

Query Match 34.1%; Score 31; DB 16; Length 35;
Best Local Similarity 35.7%; Pred. No. 5.4e+02;
Matches 5; Conservative 6; Mismatches 3; Indels 0; Gaps 0;

QY 4 VYDALNVLMMNII 17
DB 3 VYETLSIVQVNSI 16

RESULT 41
Q9PS47 PRELIMINARY; PRT; 49 AA.
AC Q9PS47;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE JARARAFIBRASE II (Fragment).
```

```
OS Bothrops jararaca (Jararaca).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Lepidosauria; Squamata; Scleroglossa; Serpentes; Colubroidea;
OC Viperidae; Crotalinae; Bothrops.
OX NCBI_TaxID=8724;
RN [1]
RP SEQUENCE.
RX MEDLINE=92397346; PubMed=1523677;
RA Maruyama M., Sugiki M., Yoshida E., Mihara H., Nakajima N.;
RT "Purification and characterization of two fibrinolytic enzymes from
RT Bothrops jararaca (Jararaca) venom."
RL Toxicon 30:853-864(1992).
DR HSSP; P15167; IATL.
DR InterPro; IPR001590; Reprolysin.
DR Pfam; PF01421; Reprolysin; 1.
DR PROSITE; PS50215; ADAM_MEPRO; 1.
SQ SEQUENCE 49 AA; 6112 MW; 88F064DC1DB2E44E CRC64;

Query Match 34.1%; Score 31; DB 13; Length 49;
Best Local Similarity 36.4%; Pred. No. 7.5e+02;
Matches 4; Conservative 5; Mismatches 2; Indels 0; Gaps 0;

QY 1 RRVYDALNVL 11
DB 30 RRRHQMVNIM 40

RESULT 42
Q8F925 PRELIMINARY; PRT; 62 AA.
AC Q8F925;
DT 01-MAR-2003 (TrEMBLrel. 23, Created)
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE Predicted transcriptional regulator, copG family.
GN LA0372.
OS Leptospira interrogans.
OC Bacteria; Spirochaetes; Spirochaetales; Leptospiraceae; Leptospira.
OX NCBI_TaxID=173;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=56601 / Serogroup Icterohaemorrhagiae / Serovar lai;
RA Ren S.;
RL Submitted (MAR-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AB011224; AAN47571.1; -
KW Complete proteome.
SQ SEQUENCE 62 AA; 7420 MW; D838FE2F7F71B70A CRC64;

Query Match 34.1%; Score 31; DB 16; Length 62;
Best Local Similarity 42.9%; Pred. No. 9.4e+02;
Matches 9; Conservative 3; Mismatches 7; Indels 2; Gaps 1;

QY 1 RRVY--DALNVLMMNIIISK 19
DB 40 RNEVRSYSYERLEALKISK 60

RESULT 43
Q8YP72 PRELIMINARY; PRT; 68 AA.
AC Q8YP72;
DT 01-MAR-2002 (TrEMBLrel. 20, Created)
DT 01-MAR-2002 (TrEMBLrel. 20, Last sequence update)
DT 01-MAR-2002 (TrEMBLrel. 20, Last annotation update)
DE Hypothetical protein Asl4328.
GN Asl4328.
OS Anabaena sp. (strain PCC 7120).
OC Bacteria; Cyanobacteria; Nostocales; Nostocaceae; Nostoc.
OX NCBI_TaxID=103690;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=21595285; PubMed=11759840;
RA Kaneko T., Nakamura Y., Wolk C.P., Kuritz T., Sasamoto S.,
```

RA Watanabe A., Iriguchi M., Ishikawa A., Kawashima K., Kimura T.,  
 RA Kishida Y., Kohara M., Matsumoto M., Matsuno A., Muraki A.,  
 RA Nakazaki N., Shimo S., Sugimoto M., Takazawa M., Yamada M.,  
 RA Yasuda M., Tabata S.;  
 RT "Complete genomic sequence of the filamentous nitrogen-fixing  
 cyanobacterium Anabaena sp. strain PCC 7120.";  
 RL DNA Res. 8:205-213(2001).  
 DR EMBL; AP003596; BAB76027.1; -;  
 KW Hypothetical protein; Complete proteome.  
 SQ SEQUENCE 68 AA; 7648 MW; 8876D5D9FFCC4B14 CRC64;

Query Match 34.1%; Score 31; DB 16; Length 68;  
 Best Local Similarity 66.7%; Pred. No. 1e+03;  
 Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 5 YDALNVLM 13  
 DB 43 YEAELELMA 51

## RESULT 44

Q97AT0 PRELIMINARY; PRT; 69 AA.  
 AC Q97AT0;  
 DT 01-OCT-2001 (TREMBLrel. 18, Created)  
 DT 01-OCT-2001 (TREMBLrel. 18, Last sequence update)  
 DT 01-MAR-2002 (TREMBLrel. 20, Last annotation update)  
 DE Hypothetical protein TV0729.  
 GN TV0729 OR TVG0736303.  
 OS Thermoplasma volcanium.  
 OC Archaea; Euryarchaeota; Thermoplasma; Thermoplasmatales;  
 OC Thermoplasmataceae; Thermoplasma.  
 OX NCBI\_TaxID=50339;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=GSS1 / DSM 4299 / JCM 9571;  
 RX MEDLINE=20570466; PubMed=11121031;  
 RA Kawashima T., Amano N., Koike H., Makino S.-I., Higuchi S.,  
 RA Kawashima-Ohya Y., Watanabe K., Yamazaki M., Kanehori K., Kawamoto T.,  
 RA Nunoshiba T., Yamamoto Y., Aramaki H., Makino K., Suzuki M.;  
 RT "Archaeal adaptation to higher temperatures revealed by genomic  
 sequence of Thermoplasma volcanium.";  
 RL Proc. Natl. Acad. Sci. U.S.A. 97:14257-14262(2000).  
 DR EMBL; AP000993; BAB59871.1; -;  
 KW Hypothetical protein; Complete proteome.  
 SQ SEQUENCE 69 AA; 8075 MW; 5018925AAB1C560E CRC64;

Query Match 34.1%; Score 31; DB 17; Length 69;  
 Best Local Similarity 31.6%; Pred. No. 1e+03;  
 Matches 6; Conservative 5; Mismatches 8; Indels 0; Gaps 0;

QY 1 RRRVYDALNVLMAMNIISK 19  
 DB 11 RRRKSDIINIKMTIVIK 29

## RESULT 45

Q96IE9 PRELIMINARY; PRT; 70 AA.  
 AC Q96IE9;  
 DT 01-DEC-2001 (TREMBLrel. 19, Created)  
 DT 01-DEC-2001 (TREMBLrel. 19, Last sequence update)  
 DT 01-OCT-2002 (TREMBLrel. 22, Last annotation update)  
 DE Hypothetical protein.  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 OX NCBI\_TaxID=9606;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Placenta;  
 RA Strausberg R.;  
 RL Submitted (MAY-2001) to the EMBL/GenBank/DBJ databases.

DR EMBL; BC007583; AAH07583.1; -;  
 DR InterPro; IPR001014; Ribosomal\_L23.  
 DR Pfam; PF00276; Ribosomal\_L23; 1.  
 DR ProDom; PD001141; Ribosomal\_L23; 1.  
 DR PROSITE; PS00050; RIBOSOMAL\_L23; 1.  
 KW Hypothetical protein.  
 SQ SEQUENCE 70 AA; 7923 MW; AC1C466548F343C1 CRC64;  
 Query Match 34.1%; Score 31; DB 4; Length 70;  
 Best Local Similarity 53.8%; Pred. No. 1.1e+03;  
 Matches 7; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

QY 5 YDALNVLMAMNII 17  
 DB 58 YDALDVANKIGII 70

Search completed: February 17, 2004, 10:56:19  
 Job time : 28.901 secs





XX DR WPI; 1998-377596/32.

XX PT Polypeptide fragments of the DP-1 transcription factor - used for

XX PT inducing apoptosis, specifically in tumour and cardiovascular cells,

XX PT e.g. for preventing restenosis

XX PS Claim 4; Page 44; 55pp; English.

XX CC Peptide H3 comprises amino acid residues 174-193 in the DEF box

XX CC (I) (see AAW30501) of transcription factor DP1. Claimed peptides

XX CC (II) (see AAW30504-07) containing one or both of 2 motifs (see

XX CC AAW30502-03) of the DEF box are capable of antagonising the

XX CC heterodimerisation of a DP protein with an E2F protein. Also

XX CC claimed are variants of these peptides, especially containing

XX CC substitutions of residues corresponding to residues 167, 169, 171

XX CC and 175 of DP-1, fusion proteins (III) comprising (I) or (II) and a

XX CC membrane translocation sequence (see AAW30508), expression vectors

XX CC encoding (I)-(III) and host cells. (I)-(III) are used

XX CC therapeutically to induce apoptosis, specifically in tumour or

XX CC cardiovascular cells, either in vivo or in vitro, e.g. for purging

XX CC bone marrow. Surgical stents comprising (I)-(III) are used to

XX CC treat or prevent restenosis in patients who have undergone

XX CC angioplasty. (I)-(III) function by inactivating the DNA-binding

XX CC activity of DP/E2F heterodimers. They are also used as research

XX CC reagents, as positive controls in assays for identifying

XX CC antagonists of DP-1/E2F dimerisation and as immunoassay agents.

XX CC Also described is the use of sequences antisense to nucleic acids

XX CC encoding (I)-(III) to control DP levels in cells, particularly by

XX CC gene therapy. When formulated with cytotoxic or cytostatic agents,

XX CC (I)-(III) enhance cell killing.

XX SQ Sequence 20 AA;

Query Match 100.0%; Score 101; DB 19; Length 20;

Best Local Similarity 100.0%; Pred. No. 2e-09;

Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 NVLMAMNIISKEKEIKWIG 20

Db 1 NVLMAMNIISKEKEIKWIG 20

RESULT 2

AAW30507

ID AAW30507 standard; Peptide; 30 AA.

XX AC AAW30507;

XX DT 26-OCT-1998 (first entry)

XX DE DP-1 transcription factor antagonist peptide H7.

XX KW DP-1; transcription factor; antagonist; E2F protein; apoptosis;

XX KW cell proliferation; cardiovascular cell; restenosis; tumour;

XX KW surgical stent; therapy.

XX OS Synthetic.

XX OS Homo sapiens.

XX FH Key Location/Qualifiers

FT Peptide 3..9

FT Peptide /note= "Claim 3"

FT Peptide 5..15

FT Peptide /note= "Claim 3"

XX PN WO9828334-A1.

XX PD 02-JUL-1998.

XX PF 22-DEC-1997; 97WO-GB03506.

XX PR 20-DEC-1996; 96GB-0026589.

XX PA (PROL-) PROLIFIX LTD.

XX PI Bandara LR, La Thangue NB;

XX DR WPI; 1998-377596/32.

XX PT Polypeptide fragments of the DP-1 transcription factor - used for

XX PT inducing apoptosis, specifically in tumour and cardiovascular cells,

XX PT e.g. for preventing restenosis

XX PS Claim 4; Page 44; 55pp; English.

XX CC Peptide H7 comprises amino acid residues 170-199 in the DEF box

XX CC (I) (see AAW30501) of transcription factor DP1. Claimed peptides

XX CC (II) (see AAW30504-07) containing one or both of 2 motifs (see

XX CC AAW30502-03) of the DEF box are capable of antagonising the

XX CC heterodimerisation of a DP protein with an E2F protein. Also

XX CC claimed are variants of these peptides, especially containing

XX CC substitutions of residues corresponding to residues 167, 169, 171

XX CC and 175 of DP-1, fusion proteins (III) comprising (I) or (II) and a

XX CC membrane translocation sequence (see AAW30508), expression vectors

XX CC encoding (I)-(III) and host cells. (I)-(III) are used

XX CC therapeutically to induce apoptosis, specifically in tumour or

XX CC cardiovascular cells, either in vivo or in vitro, e.g. for purging

XX CC bone marrow. Surgical stents comprising (I)-(III) are used to

XX CC treat or prevent restenosis in patients who have undergone

XX CC angioplasty. (I)-(III) function by inactivating the DNA-binding

XX CC activity of DP/E2F heterodimers. They are also used as research

XX CC reagents, as positive controls in assays for identifying

XX CC antagonists of DP-1/E2F dimerisation and as immunoassay agents.

XX CC Also described is the use of sequences antisense to nucleic acids

XX CC encoding (I)-(III) to control DP levels in cells, particularly by

XX CC gene therapy. When formulated with cytotoxic or cytostatic agents,

XX CC (I)-(III) enhance cell killing.

XX SQ Sequence 30 AA;

Query Match 100.0%; Score 101; DB 19; Length 30;

Best Local Similarity 100.0%; Pred. No. 3.1e-09;

Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 NVLMAMNIISKEKEIKWIG 20

Db 5 NVLMAMNIISKEKEIKWIG 24

RESULT 3

AAW30501

ID AAW30501 standard; Peptide; 37 AA.

XX AC AAW30501;

XX DT 26-OCT-1998 (first entry)

XX DE DP-1 transcription factor peptide H (DEF box).

XX KW DP-1; transcription factor; antagonist; E2F protein; apoptosis;

XX KW cell proliferation; cardiovascular cell; restenosis; tumour;

XX KW surgical stent; therapy.

XX OS Synthetic.

XX OS Homo sapiens.

XX PN WO9828334-A1.

XX PD 02-JUL-1998.

XX PF 22-DEC-1997; 97WO-GB03506.

XX PR 20-DEC-1996; 96GB-0026589.

XX PA (PROL-) PROLIFIX LTD.





PN W0953075-A2.  
 XX  
 PD 21-OCT-1999.  
 XX  
 PF 08-APR-1999; 99WO-US07638.  
 XX  
 PR 09-APR-1998; 98US-0081132.  
 XX  
 XX (DUPO ) DU PONT DE NEMOURS & CO E I.  
 XX  
 PI Klein TM, Morakinyo LO, Odell JT, Sakai H;  
 XX  
 DR WPI; 1999-633830/54.  
 XX  
 DR N-PSDB; AA234579.  
 XX  
 XX Plant-derived cell cycle regulatory proteins -  
 PT  
 XX Claim 10; Page 41; 44pp; English.  
 PS  
 XX This sequence represents 42% of the middle region of soybean cell  
 CC cycle regulatory protein DP-1, as deduced from an isolated  
 CC cDNA clone (see AA234579). The invention relates to nucleic acid  
 CC fragments (see AA234575-83) encoding plant CDC-16, DP-1, DP-2 and  
 CC E2F cell cycle regulatory proteins (see AA232159-67). It also  
 CC relates to the construction of a chimeric gene encoding all or a  
 CC portion of the cell cycle regulatory protein, in sense or antisense  
 CC orientation, where expression of the chimeric gene results in  
 CC production of altered levels of the cell cycle regulatory protein in  
 CC a transformed host cell. The nucleic acids and proteins may be  
 CC used to facilitate studies of cell cycle regulation in plants,  
 CC provide genetic tools to enhance cell growth in tissue culture,  
 CC increase gene transfer efficiency and provide more stable  
 CC transformations. The proteins may also provide targets to  
 CC facilitate design and/or identification of cell cycle regulatory  
 CC proteins that may be useful as herbicides.  
 XX  
 SQ Sequence 83 AA;  
 Query Match 68.3%; Score 69; DB 20; Length 83;  
 Best Local Similarity 70.0%; Pred. No. 0.0014;  
 Matches 14; Conservative 3; Mismatches 3; Indels 0; Gaps 0;  
 QY 1 NVLMAMNIISKEKEIKWIG 20  
 Db 15 NVLMAMNDIISXKKKIQWRG 34  
 RESULT 6  
 AAW30516  
 ID AAW30516 standard; Peptide; 19 AA.  
 XX  
 AC AAW30516;  
 XX  
 DT 26-OCT-1998 (first entry)  
 XX  
 DE DP-1 transcription factor antagonist peptide H2mt2.  
 XX  
 KW DP-1; transcription factor; E2F protein; apoptosis;  
 KW cell proliferation; cardiovascular cell; restenosis; tumour;  
 KW surgical stent; therapy.  
 XX  
 OS Synthetic.  
 OS Homo sapiens.  
 XX  
 XX Key Location/Qualifiers  
 FH Misc-difference 2 /note= "R167A mutation"  
 FT FT  
 FT Misc-difference 6 /note= "D171A mutation"  
 FT FT  
 XX W09828334-A1.  
 PN  
 XX 02-JUL-1998.  
 PD

XX 22-DEC-1997; 97WO-GB03506.  
 XX  
 PR 20-DEC-1996; 96GB-0026589.  
 XX  
 PA (PROL-) PROLIFIX LTD.  
 XX  
 PI Bandara LR, La Thangue NB;  
 XX  
 XX WPI; 1998-377596/32.  
 DR  
 XX Polypeptide fragments of the DP-1 transcription factor - used for  
 PT inducing apoptosis, specifically in tumour and cardiovascular cells,  
 PT e.g. for preventing restenosis  
 XX  
 PS Example D; Page 26; 55pp; English.  
 XX  
 CC Peptide H2mt2 is based on peptide H2 (see AAW30504) from the DEF box  
 CC (see AAW30501) of transcription factor Dp1. In H2mt2, amino acid  
 CC residues of H2 that correspond to Dp1 residues Arg167 and Asp171  
 CC are substituted by Ala residues. H2 is an antagonist of the  
 CC heterodimerisation of Dp1 with E2F. H2mt2 retains some, but not  
 CC all, of this antagonistic activity. H2 and other claimed peptides  
 CC (see AAW30504-07) from the DEF box region of Dp1 can be used to  
 CC induce apoptosis, specifically in tumour and cardiovascular cells,  
 CC e.g. for the prevention of restenosis.  
 XX  
 SQ Sequence 19 AA;  
 Query Match 50.5%; Score 51; DB 19; Length 19;  
 Best Local Similarity 100.0%; Pred. No. 0.22;  
 Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 NVLMAMNIISK 11  
 Db 9 NVLMAMNIISK 19  
 RESULT 7  
 AAW30504  
 ID AAW30504 standard; Peptide; 19 AA.  
 XX  
 AC AAW30504;  
 XX  
 DT 26-OCT-1998 (first entry)  
 XX  
 DE DP-1 transcription factor antagonist peptide H2.  
 XX  
 KW DP-1; transcription factor; antagonist; E2F protein; apoptosis;  
 KW cell proliferation; cardiovascular cell; restenosis; tumour;  
 KW surgical stent; therapy.  
 XX  
 OS Synthetic.  
 OS Homo sapiens.  
 XX  
 XX Key Location/Qualifiers  
 FH Peptide 9..18  
 FT /note= "Claim 3"  
 FT  
 XX W09828334-A1.  
 PN  
 XX 02-JUL-1998.  
 PD  
 XX 22-DEC-1997; 97WO-GB03506.  
 PF  
 XX 20-DEC-1996; 96GB-0026589.  
 PR  
 XX (PROL-) PROLIFIX LTD.  
 XX  
 PI Bandara LR, La Thangue NB;  
 XX  
 XX WPI; 1998-377596/32.  
 DR  
 XX

PT Polypeptide fragments of the DP-1 transcription factor - used for  
 PT inducing apoptosis, specifically in tumour and cardiovascular cells,  
 PT e.g. for preventing restenosis

PS Claim 4; Page 44; 55pp; English.

XX Peptide H2 comprises amino acid residues 166-184 in the DEF box  
 CC (I) (see AAW30501) of transcription factor Dp1. Claimed peptides  
 CC (II) (see AAW30504-07) containing one or both of 2 motifs (see  
 CC AAW30502-03) of the DEF box are capable of antagonising the  
 CC heterodimerisation of a Dp protein with an E2F protein. Also  
 CC claimed are variants of these peptides, especially containing  
 CC substitutions of residues corresponding to residues 167, 169, 171  
 CC and 175 of DP-1, fusion proteins (III) comprising (I) or (II) and a  
 CC membrane translocation sequence (see AAW30508), expression vectors  
 CC encoding (I)-(III) and host cells. (I)-(III) are used  
 CC therapeutically to induce apoptosis, specifically in tumour or  
 CC cardiovascular cells, either in vivo or in vitro, e.g. for purging  
 CC bone marrow. Surgical stents comprising (I)-(III) are used to  
 CC treat or prevent restenosis in patients who have undergone  
 CC angioplasty. (I)-(III) function by inactivating the DNA-binding  
 CC activity of DP/E2F heterodimers. They are also used as research  
 CC reagents, as positive controls in assays for identifying  
 CC antagonists of DP-1/E2F dimerisation and as immunoassay agents.  
 CC Also described is the use of sequences antisense to nucleic acids  
 CC encoding (I)-(III) to control DP levels in cells, particularly by  
 CC gene therapy. When formulated with cytotoxic or cytostatic agents,  
 CC (I)-(III) enhance cell killing.

XX Sequence 19 AA;

Query Match 50.5%; Score 51; DB 19; Length 19;

Best Local Similarity 100.0%; Pred. No. 0.22;

Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 NVLMAMNIIISK 11

DB 9 NVLMAMNIIISK 19

RESULT 8

AAU72601  
 ID AAU72601 standard; Peptide; 35 AA.

XX AAU72601;

DT 26-FEB-2002 (first entry)

XX DEF domain consensus sequence.

XX Cell cycle protein; CCP; cell cycle regulation; herbicide;  
 KW plant growth regulator; plant development; abiotic stress; biotic stress;  
 KW nutrient deprivation; pathogen attack; crop yield; motif.

OS Synthetic.

XX WO200185946-A2.

PN 15-NOV-2001.

PD 14-MAY-2001; 2001WO-IB01307.

XX 12-MAY-2000; 2000US-204045P.

XX (CROP-) CROPDISEIGN NV.

XX Inze D, Boudolf V, De Veylder L, Acosta JAT, Magyar Z;

PI WPI; 2002-062249/08.

XX New cell cycle protein and nucleic acid molecule encoding it useful for  
 PT regulating cell cycle progression in plants and for identifying  
 PT modulators which are useful as herbicides or plant growth regulators -

XX Disclosure; Page 25; 316pp; English.

XX The invention relates to a novel cell cycle protein (CCP) and the  
 CC polynucleotides encoding them. CCP is useful for identifying a compound  
 CC which modulates the activity of the polypeptide and which binds to the  
 CC polypeptide and an anti-CCP antibody is useful for detecting the presence  
 CC of CCP in a sample. A CCP modulator is useful for modulating the cell  
 CC cycle or growth of a plant such as Arabidopsis thaliana, rice, wheat,  
 CC maize, tomato, alfalfa, oilseed rape, soybean, sunflower and canola.  
 CC CCP nucleic acid and polypeptide molecules are useful as modulating  
 CC agents in regulating cell cycle progression in plants. CCP is useful to  
 CC treat disorders characterised by insufficient or excessive production of  
 CC CCP protein or production of CCP protein forms which have decreased or  
 CC aberrant activity. Compounds that bind to or modulate the activity  
 CC of CCP polypeptide are useful as herbicides or plant growth regulators.  
 CC The polynucleotide is useful for modifying cell fate, plant development,  
 CC plant morphology, biochemistry and/or physiology, the length of the G1,  
 CC S, G2 and/or M phase of the cell cycle of a plant, initiation, promotion,  
 CC stimulation or enhancement of cell division, DNA replication, seed set,  
 CC seed size, seed development, tuber, fruit, leaf formation, shoot and root  
 CC initiation and/or development, nodule function, dwarfism in plants,  
 CC senescence, tolerance or resistance to stress. CCP, the polynucleotide  
 CC and the anti-CCP antibody are useful in agriculture to modulate the  
 CC protein levels or activity of a protein involved in the cell cycle due  
 CC to environmental conditions, including abiotic stress such as  
 CC cold, nutrient deprivation, heat, drought, salt stress, or biotic  
 CC stress such as pathogen attack, to modulate e.g. enhance crop yields,  
 CC and attenuate plant architecture, plant quality traits, plant  
 CC reproduction and seed development, endoreduplication in storage cells,  
 CC storage tissues and/or storage organs of plants or its parts. CCP is  
 CC useful as an immunogen to generate antibodies. CCP protein is useful to  
 CC screen for naturally occurring CCP substrates. The polynucleotide is  
 CC useful for expressing CCP protein, to detect CCP mRNA, or a genetic  
 CC lesion in a CCP gene and to modulate CCP activity. The present sequence  
 CC represents a motif which may be found in a CCP protein of the invention.

XX Sequence 35 AA;

Query Match 50.5%; Score 51; DB 23; Length 35;

Best Local Similarity 55.0%; Pred. No. 0.43;

Matches 11; Conservative 0; Mismatches 9; Indels 0; Gaps 0;

QY 1 NVLMAMNIIISKKEIKWIG 20

DB 14 NVYMAXXXIXXXKEIXWKG 33

RESULT 9

AAW30510

ID AAW30510 standard; Peptide; 15 AA.

XX AAW30510;

XX 26-OCT-1998 (first entry)

XX DP-1 transcription factor peptide H4.

XX DP-1; transcription factor; antagonist; E2F protein; apoptosis;  
 KW cell proliferation; cardiovascular cell; restenosis; tumour;  
 KW surgical stent; therapy.

OS Synthetic.

OS Homo sapiens.

XX WO9828334-A1.

PN 02-JUL-1998.

XX 22-DEC-1997; 97WO-GB03506.

XX 20-DEC-1996; 96GB-0026589.

XX

```

PA (PROL-) PROLIFIX LTD.
XX
PI Bandara LR, La Thangue NB;
XX
DR WPI; 1998-377596/32.
XX
PT Polypeptide fragments of the DP-1 transcription factor - used for
PT inducing apoptosis, specifically in tumour and cardiovascular cells,
PT e.g. for preventing restenosis
XX
PS Example C; Page 41; 55pp; English.
XX
CC Peptide H4 comprises amino acid residues 185-199 in the DEF box
CC region (see AA030501) of transcription factor Dp1. Unlike claimed
CC peptides (see AA030504-07) that contain one or both of 2 motifs (see
CC AA030502-03) of the Dp1 DEF box, peptide H4 is not capable of
CC antagonising the heterodimerisation of a DP protein with an E2F
CC protein. The claimed peptides, their variants and fusion proteins
CC can be used to induce apoptosis, specifically in tumour and
CC cardiovascular cells, e.g. to prevent restenosis.
XX
SQ Sequence 15 AA;
Query Match 49.5%; Score 50; DB 19; Length 15;
Best Local Similarity 100.0%; Pred. No. 0.24;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 12 EKKEIKWIG 20
DB 1 EKKEIKWIG 9
|||||||
|

RESULT 10
ABB03433
ID ABB03433 standard; Protein; 29 AA.
AC ABB03433;
XX
DT 08-JAN-2002 (first entry)
XX
DE Human musculoskeletal system related polypeptide SEQ ID NO 1380.
XX
KW Cytostatic; immunosuppressive; nootropic; neuroprotective; antiviral;
KW antiallergic; hepatotropic; antidiabetic; antiinflammatory; antiulcer;
KW vulnary; anticoagulant; antibacterial; antifungal; antiparasitic;
KW cardiant; gene therapy; cancer; immune disorder; cardiovascular disorder;
KW neurological disease; infection; human; secreted protein;
KW musculoskeletal system.
XX
OS Homo sapiens.
XX
PN WO200155367-A1.
XX
PD 02-AUG-2001.
XX
PF 17-JAN-2001; 2001WO-US01338.
XX
PR 31-JAN-2000; 2000US-0179065.
PR 04-FEB-2000; 2000US-0180628.
PR 24-FEB-2000; 2000US-0184664.
PR 02-MAR-2000; 2000US-0186350.
PR 16-MAR-2000; 2000US-0189874.
PR 17-MAR-2000; 2000US-0190076.
PR 18-APR-2000; 2000US-0198123.
PR 19-MAY-2000; 2000US-0205515.
PR 07-JUN-2000; 2000US-0209467.
PR 28-JUN-2000; 2000US-0214886.
PR 30-JUN-2000; 2000US-0215135.
PR 07-JUL-2000; 2000US-0216647.
PR 07-JUL-2000; 2000US-0216880.
PR 11-JUL-2000; 2000US-0217487.
PR 11-JUL-2000; 2000US-0217496.
PR 14-JUL-2000; 2000US-0218290.
PR 26-JUL-2000; 2000US-0220963.
PR 26-JUL-2000; 2000US-0220964.
PR 14-AUG-2000; 2000US-0224518.
PR 14-AUG-2000; 2000US-0224519.
PR 14-AUG-2000; 2000US-0225213.
PR 14-AUG-2000; 2000US-0225214.
PR 14-AUG-2000; 2000US-0225266.
PR 14-AUG-2000; 2000US-0225267.
PR 14-AUG-2000; 2000US-0225268.
PR 14-AUG-2000; 2000US-0225270.
PR 14-AUG-2000; 2000US-0225447.
PR 14-AUG-2000; 2000US-0225757.
PR 14-AUG-2000; 2000US-0225758.
PR 18-AUG-2000; 2000US-0225759.
PR 22-AUG-2000; 2000US-0226279.
PR 22-AUG-2000; 2000US-0226681.
PR 22-AUG-2000; 2000US-0228668.
PR 22-AUG-2000; 2000US-0228669.
PR 23-AUG-2000; 2000US-0227182.
PR 30-AUG-2000; 2000US-0228924.
PR 01-SEP-2000; 2000US-0229287.
PR 01-SEP-2000; 2000US-0229143.
PR 01-SEP-2000; 2000US-0229344.
PR 01-SEP-2000; 2000US-0229345.
PR 05-SEP-2000; 2000US-0229509.
PR 05-SEP-2000; 2000US-0229513.
PR 06-SEP-2000; 2000US-0230437.
PR 06-SEP-2000; 2000US-0230438.
PR 08-SEP-2000; 2000US-0231242.
PR 08-SEP-2000; 2000US-0231243.
PR 08-SEP-2000; 2000US-0231244.
PR 08-SEP-2000; 2000US-0231413.
PR 08-SEP-2000; 2000US-0231414.
PR 08-SEP-2000; 2000US-0232080.
PR 08-SEP-2000; 2000US-0232081.
PR 12-SEP-2000; 2000US-0231968.
PR 14-SEP-2000; 2000US-0232397.
PR 14-SEP-2000; 2000US-0232398.
PR 14-SEP-2000; 2000US-0232399.
PR 14-SEP-2000; 2000US-0232400.
PR 14-SEP-2000; 2000US-0232401.
PR 14-SEP-2000; 2000US-0233063.
PR 14-SEP-2000; 2000US-0233064.
PR 21-SEP-2000; 2000US-0233065.
PR 21-SEP-2000; 2000US-0234223.
PR 21-SEP-2000; 2000US-0234274.
PR 25-SEP-2000; 2000US-0234997.
PR 25-SEP-2000; 2000US-0234998.
PR 26-SEP-2000; 2000US-0235484.
PR 27-SEP-2000; 2000US-0235834.
PR 29-SEP-2000; 2000US-0235836.
PR 29-SEP-2000; 2000US-0236327.
PR 29-SEP-2000; 2000US-0236367.
PR 29-SEP-2000; 2000US-0236368.
PR 29-SEP-2000; 2000US-0236369.
PR 29-SEP-2000; 2000US-0236370.
PR 02-OCT-2000; 2000US-0236802.
PR 02-OCT-2000; 2000US-0237037.
PR 02-OCT-2000; 2000US-0237038.
PR 02-OCT-2000; 2000US-0237039.
PR 02-OCT-2000; 2000US-0237040.
PR 13-OCT-2000; 2000US-0239935.
PR 13-OCT-2000; 2000US-0239937.
PR 20-OCT-2000; 2000US-0240960.
PR 20-OCT-2000; 2000US-0241221.
PR 20-OCT-2000; 2000US-0241785.
PR 20-OCT-2000; 2000US-0241786.
PR 20-OCT-2000; 2000US-0241787.
PR 20-OCT-2000; 2000US-0241808.
PR 20-OCT-2000; 2000US-0241809.
PR 20-OCT-2000; 2000US-0241826.
PR 01-NOV-2000; 2000US-0244617.
PR 08-NOV-2000; 2000US-0246474.

```

PR 08-NOV-2000; 2000US-0246475.  
 PR 08-NOV-2000; 2000US-0246476.  
 PR 08-NOV-2000; 2000US-0246477.  
 PR 08-NOV-2000; 2000US-0246478.  
 PR 08-NOV-2000; 2000US-0246523.  
 PR 08-NOV-2000; 2000US-0246524.  
 PR 08-NOV-2000; 2000US-0246525.  
 PR 08-NOV-2000; 2000US-0246526.  
 PR 08-NOV-2000; 2000US-0246527.  
 PR 08-NOV-2000; 2000US-0246528.  
 PR 08-NOV-2000; 2000US-0246532.  
 PR 08-NOV-2000; 2000US-0246609.  
 PR 08-NOV-2000; 2000US-0246610.  
 PR 08-NOV-2000; 2000US-0246611.  
 PR 08-NOV-2000; 2000US-0246613.  
 PR 17-NOV-2000; 2000US-0249207.  
 PR 17-NOV-2000; 2000US-0249208.  
 PR 17-NOV-2000; 2000US-0249209.  
 PR 17-NOV-2000; 2000US-0249210.  
 PR 17-NOV-2000; 2000US-0249211.  
 PR 17-NOV-2000; 2000US-0249212.  
 PR 17-NOV-2000; 2000US-0249213.  
 PR 17-NOV-2000; 2000US-0249214.  
 PR 17-NOV-2000; 2000US-0249215.  
 PR 17-NOV-2000; 2000US-0249216.  
 PR 17-NOV-2000; 2000US-0249217.  
 PR 17-NOV-2000; 2000US-0249218.  
 PR 17-NOV-2000; 2000US-0249219.  
 PR 17-NOV-2000; 2000US-0249299.  
 PR 01-DEC-2000; 2000US-0249300.  
 PR 01-DEC-2000; 2000US-0250160.  
 PR 03-DEC-2000; 2000US-0250391.  
 PR 03-DEC-2000; 2000US-0251030.  
 PR 05-DEC-2000; 2000US-0251988.  
 PR 05-DEC-2000; 2000US-0256719.  
 PR 06-DEC-2000; 2000US-0251479.  
 PR 06-DEC-2000; 2000US-0251856.  
 PR 08-DEC-2000; 2000US-0251868.  
 PR 08-DEC-2000; 2000US-0251869.  
 PR 08-DEC-2000; 2000US-0251989.  
 PR 08-DEC-2000; 2000US-0251990.  
 PR 11-DEC-2000; 2000US-0254097.  
 PR 05-JAN-2001; 2001US-0259678.  
 (HUMA-) HUMAN GENOME SCI INC.

XX Rosen CA, Barash SC, Ruben SM;  
 PI WPI; 2001-451937/48.  
 XX N-PSDB; AAL35015.

XX Isolated polypeptide for treating, preventing and/ or prognosing  
 PT disorders related to the musculoskeletal system including  
 PT musculoskeletal cancers and also for testing and detection e.g.  
 PT diagnosis -

XX Claim 11; SEQ ID NO 1380; 781pp + Sequence Listing; English.

XX The invention relates to novel genes (AAL34669-AAL37666) and proteins  
 CC (ABB03087-ABB04109) associated with the musculoskeletal system useful  
 CC for preventing, treating or ameliorating medical conditions e.g. by  
 CC protein or gene therapy. The genes are isolated from a range of human  
 CC tissues disclosed in the specification. The nucleic acids, proteins,  
 CC antibodies and (ant)agonists are useful in the diagnosis, treatment  
 CC and prevention of: (a) cancer, e.g. breast and ovarian cancer and  
 CC other cancers of the adrenal gland, bone, bone marrow, breast,  
 CC gastrointestinal tract, liver, lung, or urogenital; (b) immune  
 CC disorders e.g. Addison's disease, allergies, autoimmune haemolytic  
 CC anaemia, autoimmune thyroiditis, diabetes mellitus, Crohn's disease,

CC multiple sclerosis, rheumatoid arthritis and ulcerative colitis;  
 CC (c) cardiovascular disorders such as myocardial ischaemias; (d) wound  
 CC healing; (e) neurological diseases e.g. cerebral anoxia and epilepsy;  
 CC and (f) infectious diseases such as viral, bacterial, fungal and  
 CC parasitic infections.  
 CC Note: The sequence data for this patent did not form part of the  
 CC printed specification, but was obtained in electronic format directly  
 CC from WIPO at ftp.wipo.int/pub/published\_pct\_sequences.  
 XX  
 XX Sequence 29 AA;

Query Match 49.5%; Score 50; DB 22; Length 29;  
 Best Local Similarity 66.7%; Pred. No. 0.51;  
 Matches 10; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY 6 MNIISKEKEIKWIG 20  
 | | | | | | | | | |  
 Db 13 MKIFSKEKKIGWPG 27

## RESULT 11

ABU12727  
 ID ABU12727 standard; Protein; 29 AA.

XX AC ABU12727;

DT 26-FEB-2003 (first entry)

XX DE Novel human musculoskeletal system antigen #347.

KW Musculoskeletal system antigen; cancer; metastasis;  
 KW re-vascularisation; thrombosis; arteriosclerosis; mineral content;  
 KW cardiovascular condition; wound; injury; burn; angiogenesis; ulcer;  
 KW post-operative tissue repair; limb regeneration; neuronal growth;  
 KW neurodegenerative disorder; Alzheimer's disease; Parkinson's disease;  
 KW AIDS-related complex; chondrocyte growth; bone regeneration;  
 KW periodontal regeneration; tissue transport; bone graft; skin aging;  
 KW keratinocyte growth; hair loss; melanocyte growth; cell proliferation;  
 KW cell growth; organ transplant; cell differentiation; body height;  
 KW weight; hair colour; eye colour; skin; percentage of adipose tissue;  
 KW pigmentation; cosmetic surgery; metabolism; biorhythm; cardiac rhythm;  
 KW depression; tendency for violence; pain; reproductive capability;  
 KW hormone level; endocrine level; appetite; libido; memory; stress;  
 KW storage capability; fat content; lipid content; protein content;  
 KW carbohydrate content; vitamin content; cofactor content;  
 KW nutritional component.

XX Homo sapiens.

XX US2002147140-A1.

XX PD 10-OCT-2002.

XX PF 17-JAN-2001; 2001US-0764877.

XX 31-JAN-2000; 2000US-179065P.  
 PR 04-FEB-2000; 2000US-180628P.  
 PR 28-JUN-2000; 2000US-214896P.  
 PR 07-JUL-2000; 2000US-216647P.  
 PR 07-JUL-2000; 2000US-216880P.  
 PR 11-JUL-2000; 2000US-217487P.  
 PR 11-JUL-2000; 2000US-217496P.  
 PR 14-JUL-2000; 2000US-218290P.  
 PR 26-JUL-2000; 2000US-220963P.  
 PR 26-JUL-2000; 2000US-220964P.  
 PR 14-AUG-2000; 2000US-224518P.  
 PR 14-AUG-2000; 2000US-224519P.  
 PR 14-AUG-2000; 2000US-225267P.  
 PR 14-AUG-2000; 2000US-225268P.  
 PR 14-AUG-2000; 2000US-225270P.  
 PR 14-AUG-2000; 2000US-225447P.  
 PR 14-AUG-2000; 2000US-225757P.  
 PR 14-AUG-2000; 2000US-225759P.

PR 22-AUG-2000; 2000US-226868P.  
PR 30-AUG-2000; 2000US-228924P.  
PR 01-SEP-2000; 2000US-229287P.  
PR 01-SEP-2000; 2000US-229343P.  
PR 01-SEP-2000; 2000US-229344P.  
PR 01-SEP-2000; 2000US-229345P.  
PR 05-SEP-2000; 2000US-229509P.  
PR 05-SEP-2000; 2000US-229513P.  
PR 08-SEP-2000; 2000US-231413P.  
PR 21-SEP-2000; 2000US-234223P.  
PR 21-SEP-2000; 2000US-234274P.  
PR 25-SEP-2000; 2000US-234997P.  
PR 27-SEP-2000; 2000US-235834P.  
PR 28-SEP-2000; 2000US-236327P.  
PR 29-SEP-2000; 2000US-236367P.  
PR 29-SEP-2000; 2000US-236368P.  
PR 29-SEP-2000; 2000US-236369P.  
PR 29-SEP-2000; 2000US-236370P.  
PR 02-OCT-2000; 2000US-236802P.  
PR 02-OCT-2000; 2000US-237037P.  
PR 02-OCT-2000; 2000US-237038P.  
PR 02-OCT-2000; 2000US-237039P.  
PR 02-OCT-2000; 2000US-237040P.  
PR 13-OCT-2000; 2000US-239935P.  
PR 20-OCT-2000; 2000US-240960P.  
PR 20-OCT-2000; 2000US-241785P.  
PR 20-OCT-2000; 2000US-241809P.  
PR 01-NOV-2000; 2000US-244617P.  
PR 17-NOV-2000; 2000US-249299P.  
PR 08-DEC-2000; 2000US-251856P.  
PR 08-DEC-2000; 2000US-251868P.  
PR 08-DEC-2000; 2000US-251869P.  
XX (ROSE/) ROSEN C A.  
PA (RUBE/) RUBEN S M.  
PA (BARA/) BARASH S C.  
XX Rosen CA, Ruben SM, Barash SC;  
XX WPI; 2003-120199/12.  
XX N-PSDB; ABX59003.  
XX Isolated nucleic acid molecules encoding musculoskeletal system  
XX associated polypeptides, useful for detecting disorders, e.g. cancer -  
XX Claim 11; SEQ ID NO 1380; 321pp; English.  
XX The invention describes an isolated nucleic acid molecule comprising a  
XX sequence encoding musculoskeletal system associated polypeptides useful  
XX for detecting disorders, e.g., cancer or cancer metastases, in animals  
XX or humans. The nucleic acid; stimulates re-vascularisation of ischaemic  
XX tissues associated with conditions such as thrombosis, arteriosclerosis,  
XX and other cardiovascular conditions; treats wounds due to injuries,  
XX burns, post-operative tissue repair, and ulcers; stimulates angiogenesis  
XX and limb regeneration; stimulates neuronal growth; can treat and prevent  
XX neuronal damage occurring in certain disorders or neurodegenerative  
XX conditions, such as, Alzheimer's disease, Parkinson's disease, and  
XX AIDS-related complex; stimulates chondrocyte growth, thus they can be  
XX used to enhance bone and periodontal regeneration and aid in tissue  
XX transports or bone grafts; prevents skin aging due to sunburn by  
XX stimulating keratinocyte growth; prevents hair loss, since FGF family  
XX members activate hair-forming cells and promotes melanocyte growth;  
XX stimulates growth and differentiation of hematopoietic cells and bone  
XX marrow cells when used in combination with other cytokines; maintains  
XX organs before transplantation or for supporting cell culture of primary  
XX tissues; induces tissue of mesodermal origin to differentiate in early  
XX embryonic; increases or decreases the differentiation or proliferation of  
XX embryonic stem cells, besides, haematopoietic lineage; modulates  
XX mammalian characteristics, such as, body height, weight, hair colour, eye  
XX colour, skin, percentage of adipose tissue, pigmentation, size, and shape  
XX (e.g., cosmetic surgery); modulates mammalian metabolism; changes  
XX mammal's metal state or physical state by influencing biorhythms,  
XX cardiac rhythms, depression, tendency for violence, tolerance for pain,

CC reproductive capabilities, hormonal or endocrine levels, appetite,  
CC libido, memory, or stress; increases or decreases storage capabilities,  
CC fat content, lipid, protein, carbohydrate, vitamins, minerals, cofactors  
CC or other nutritional components. This is the amino acid sequence of a  
CC novel human musculoskeletal system antigen.  
CC Note: The sequence data for this patent did not form part of the  
CC printed specification, but was obtained in electronic format directly  
CC from the US patent office at  
CC ftp.segdata.uspto.gov/sequence.html?docID=20020147140.  
XX  
XX Sequence 29 AA;  
SQ  
Query Match 49.5%; Score 50; DB 24; Length 29;  
Best Local Similarity 66.7%; Pred. NO. 0.51;  
Matches 10; Conservative 1; Mismatches 4; Indels 0; Gaps 0;  
QY 6 MNIISKEKKEIKWIG 20  
DB 13 MKIFSKEKKKIGWPG 27  
RESULT 12  
AAU72561  
ID AAU72561 standard; Protein; 93 AA.  
XX  
AC AAU72561;  
XX  
DT 26-FEB-2002 (first entry)  
XX  
DE Arabidopsis cell cycle protein AtDpa 121-293.  
XX  
KW Cell cycle protein; CCP; cell cycle regulation; herbicide;  
KW plant growth regulator; plant development; abiotic stress; biotic stress;  
KW nutrient deprivation; pathogen attack; crop yield; immunogen; mutant;  
KW mutuin.  
XX  
XX Arabidopsis thaliana.  
OS Synthetic.  
XX  
XX WO200185946-A2.  
PN  
XX 15-NOV-2001.  
PD  
XX  
XX 14-MAY-2001; 2001WO-IB01307.  
PF  
XX  
XX 12-MAY-2000; 2000US-204045P.  
PR  
XX  
XX (CROP-) CROPDESIGN NV.  
PA  
XX Inze D, Boudolf V, De Veylder L, Acosta JAT, Magyar Z;  
PI WPI; 2002-062249/08.  
XX  
XX New cell cycle protein and nucleic acid molecule encoding it useful for  
XX regulating cell cycle progression in plants and for identifying  
XX modulators which are useful as herbicides or plant growth regulators -  
XX Example 10; Page 285; 316pp; English.  
XX  
XX The invention relates to a novel cell cycle protein (CCP) and the  
XX polynucleotides encoding them. CCP is useful for identifying a compound  
XX which modulates the activity of the polypeptide and which binds to the  
XX polypeptide and an anti-CCP antibody is useful for detecting the presence  
XX of CCP in a sample. A CCP modulator is useful for modulating the cell  
XX cycle or growth of a plant such as Arabidopsis thaliana, rice, wheat,  
XX maize, tomato, alfalfa, oilseed rape, soybean, sunflower and canola.  
XX CCP nucleic acid and polypeptide molecules are useful as modulating  
XX agents in regulating cell cycle progression in plants. CCP is useful to  
XX treat disorders characterised by insufficient or excessive production of  
XX CCP protein or production of CCP protein forms which have decreased or  
XX aberrant activity. Compounds that bind to or modulate the activity  
XX of CCP polypeptide are useful as herbicides or plant growth regulators.  
XX The polynucleotide is useful for modifying cell fate, plant development,

CC plant morphology, biochemistry and/or physiology, the length of the G1,  
 CC S, G2 and/or M phase of the cell cycle of a plant, initiation, promotion,  
 CC stimulation or enhancement of cell division, DNA replication, seed set,  
 CC seed size, seed development, tuber, fruit, leaf formation, shoot and root  
 CC initiation and/or development, nodule function, dwarfism in plants,  
 CC senescence, tolerance or resistance to stress, CCP, the polynucleotide  
 CC and the anti-CCP antibody are useful in agriculture to modulate the  
 CC protein levels or activity of a protein involved in the cell cycle due  
 CC to environmental conditions, including abiotic stress such as  
 CC cold, nutrient deprivation, heat, drought, salt stress, or biotic  
 CC stress such as pathogen attack, to modulate e.g. enhance crop yields,  
 CC and attenuate plant architecture, plant quality traits, plant  
 CC reproduction and seed development, endoreduplication in storage cells,  
 CC storage tissues and/or storage organs of plants or its parts. CCP is  
 CC useful as an immunogen to generate antibodies. CCP protein is useful to  
 CC screen for naturally occurring CCP substrates. The polynucleotide is  
 CC useful for expressing CCP protein, to detect CCP mRNA, or a genetic  
 CC lesion in a CCP gene and to modulate CCP activity. The present sequence  
 CC represents a deletion mutant of a CCP protein of the invention.

SQ Sequence 93 AA;

Query Match 47.5%; Score 48; DB 23; Length 93;

Best Local Similarity 61.5%; Pred. No. 4;  
 Matches 8; Conservative 4; Mismatches 1; Indels 0; Gaps 0;

QY 8 IISKEKKEIKWIG 20

DB 1 IARDKKEIKWK 13

RESULT 13

ID AAW30515  
 XX AAW30515 standard; Peptide; 19 AA.

AC AAW30515;

DT 26-OCT-1998 (first entry)

DE DP-1 transcription factor peptide H2mt1.

XX DP-1; transcription factor; antagonist; E2F protein; apoptosis;  
 KW cell proliferation; cardiovascular cell; restenosis; tumour;  
 KW surgical stent; therapy.

XX Synthetic.

OS Homo sapiens.

XX Key Location/Qualifiers

FT Misc-difference 4

FT /note= "V169A mutation"

FT Misc-difference 10

FT /note= "V175A mutation"

XX WO9828334-A1.

XX 02-JUL-1998.

PF 22-DEC-1997; 97WO-GB03506.

XX 20-DEC-1996; 96GB-0026589.

XX (PROL-) PROLIFIX LTD.

PI Bandara LR, La Thangue NB;

XX WPI; 1998-377596/32.

XX Polypeptide fragments of the DP-1 transcription factor - used for  
 PT inducing apoptosis, specifically in tumour and cardiovascular cells,  
 PT e.g. for preventing restenosis

XX Example D; Page 26; 55pp; English.

XX Peptide H2mt1 is based on peptide H2 (see AAW30504) from the DEF box  
 CC (see AAW30501) of transcription factor DP1. The H2mt1 peptide, in  
 CC which H2 residues corresponding to DP1 residues Val169 and Val175  
 CC are substituted by Ala residues, behaves in a similar fashion to  
 CC the wild-type H2 peptide in its ability to inactivate E2F site DNA  
 CC binding activity in D9 EC cell extracts. H2 is an antagonist of the  
 CC heterodimerisation of DP1 with E2F. Thus, the Val-169 and Val-175  
 CC residues of H2 play little role in this activity. H2 and other  
 CC claimed peptides (see AAW30504-07) from the DEF box region of DP1 can  
 CC be used to induce apoptosis, specifically in tumour and  
 CC cardiovascular cells, e.g. for the prevention of restenosis.

XX Sequence 19 AA;

Query Match 46.5%; Score 47; DB 19; Length 19;

Best Local Similarity 90.9%; Pred. No. 0.96;  
 Matches 10; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 NVLMAMNIISK 11

DB 9 NALMAMNIISK 19

RESULT 14

ID AAW30506  
 XX AAW30506 standard; Peptide; 16 AA.

AC AAW30506;

DT 26-OCT-1998 (first entry)

DE DP-1 transcription factor antagonist peptide H5.

XX DP-1; transcription factor; antagonist; E2F protein; apoptosis;  
 KW cell proliferation; cardiovascular cell; restenosis; tumour;  
 KW surgical stent; therapy.

XX Synthetic.

OS Homo sapiens.

XX Key Location/Qualifiers

FT Peptide 5..11

FT /note= "Claim 3"

FT Peptide 7..16

FT /note= "Claim 3"

XX WO9828334-A1.

XX 02-JUL-1998.

PF 22-DEC-1997; 97WO-GB03506.

XX 20-DEC-1996; 96GB-0026589.

XX (PROL-) PROLIFIX LTD.

PI Bandara LR, La Thangue NB;

XX WPI; 1998-377596/32.

XX Polypeptide fragments of the DP-1 transcription factor - used for  
 PT inducing apoptosis, specifically in tumour and cardiovascular cells,  
 PT e.g. for preventing restenosis

XX Claim 4; Page 44; 55pp; English.

XX Peptide H5 comprises amino acid residues 168-183 in the DEF box  
 CC (I) (see AAW30501) of transcription factor DP1. Claimed peptides  
 CC (II) (see AAW30504-07) containing one or both of 2 motifs (see  
 CC AAW30502-03) of the DEF box are capable of antagonising the  
 CC heterodimerisation of a DP protein with an E2F protein. Also  
 CC claimed are variants of these peptides, especially containing

CC substitutions of residues corresponding to residues 167, 169, 171  
 CC and 175 of DP-1, fusion proteins (III) comprising (I) or (II) and a  
 CC membrane translocation sequence (see AAW30508), expression vectors  
 CC encoding (I)-(III) and host cells (I)-(III) are used  
 CC therapeutically to induce apoptosis, specifically in tumour or  
 CC cardiovascular cells, either in vivo or in vitro, e.g. for purging  
 CC bone marrow. Surgical stents comprising (I)-(III) are used to  
 CC treat or prevent restenosis in patients who have undergone  
 CC angioplasty. (I)-(III) function by inactivating the DNA-binding  
 CC activity of DP/E2F heterodimers. They are also used as research  
 CC reagents, as positive controls in assays for identifying  
 CC antagonists of DP-1/E2F dimerisation and as immunoassay agents.  
 CC Also described is the use of sequences antisense to nucleic acids  
 CC encoding (I)-(III) to control DP levels in cells, particularly by  
 CC gene therapy. When formulated with cytotoxic or cytostatic agents,  
 CC (I)-(III) enhance cell killing.

SX Sequence 16 AA;  
 Query Match 45.5%; Score 46; DB 19; Length 16;  
 Best Local Similarity 100.0%; Pred. No. 1.1;  
 Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 NVLMAMNIIS 10  
 |||||  
 Db 7 NVLMAMNIIS 16

RESULT 15  
 AAW57051  
 ID AAW57051 standard; peptide; 28 AA.

AC AAW57051;  
 XX 28-AUG-1998 (first entry)  
 DE E2F activity inhibiting compound Ib-1.  
 XX E2F activity; inhibitor; treatment; tumour; arteriosclerosis.  
 KW Synthetic.  
 OS Homo sapiens.

XX Key Location/Qualifiers  
 FH Modified-site 1 /note= "N-terminal acetyl"  
 FT Modified-site 28 /note= "C-terminal amide"  
 FT

XX WO9814474-A1.  
 PN 09-APR-1998.  
 XX 26-SEP-1997; 97WO-JP03442.  
 XX 30-SEP-1996; 96JP-0259432.  
 XX (KYOW ) KYOWA HAKKO KOGYO KK.  
 PA Mizukami T, Shibata K, Yamasaki M, Yoshida T;  
 PI WPI; 1998-240020/21.

XX E2F activity inhibitors - for treatment and prevention of tumours  
 PT and arteriosclerosis

PS Example 3; Page 27; 52pp; Japanese.

XX This represents a compound that can inhibit E2F activity. The compound  
 CC is of the formula R1 - A - R2 where R1 is an optionally substituted  
 CC alkanoyl, allyl, hetero-arylcabonyl, alkoxycabonyl, aryloxy carbonyl,  
 CC hetero-aryloxy carbonyl, or H, R2 is OH, or optionally substituted alkoxy  
 CC or amino, and A is an E2F family dimer forming region or DNA binding

CC region, of at least 12 consecutive amino acids. Compounds of this formula  
 CC can be used to inhibit E2F activity, and are useful in the treatment and  
 CC prevention of tumours and arteriosclerosis.

SX Sequence 28 AA;  
 Query Match 45.5%; Score 46; DB 19; Length 28;  
 Best Local Similarity 100.0%; Pred. No. 2.2;  
 Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 NVLMAMNIIS 10  
 |||||  
 Db 19 NVLMAMNIIS 28

RESULT 16  
 AAW57055  
 ID AAW57055 standard; peptide; 28 AA.

AC AAW57055;  
 XX 28-AUG-1998 (first entry)  
 DE E2F activity inhibiting compound Ib-3.  
 XX E2F activity; inhibitor; treatment; tumour; arteriosclerosis.  
 KW Synthetic.  
 OS Homo sapiens.

XX Key Location/Qualifiers  
 FH Modified-site 1 /note= "N-terminal lauroyl"  
 FT Modified-site 28 /note= "C-terminal amide"  
 FT

XX WO9814474-A1.  
 PN 09-APR-1998.  
 XX 26-SEP-1997; 97WO-JP03442.  
 XX 30-SEP-1996; 96JP-0259432.  
 XX (KYOW ) KYOWA HAKKO KOGYO KK.  
 PA Mizukami T, Shibata K, Yamasaki M, Yoshida T;  
 PI WPI; 1998-240020/21.

XX E2F activity inhibitors - for treatment and prevention of tumours  
 PT and arteriosclerosis

PS Example 7; Page 33; 52pp; Japanese.

XX This represents a compound that can inhibit E2F activity. The compound  
 CC is of the formula R1 - A - R2 where R1 is an optionally substituted  
 CC alkanoyl, allyl, hetero-arylcabonyl, alkoxycabonyl, aryloxy carbonyl,  
 CC hetero-aryloxy carbonyl, or H, R2 is OH, or optionally substituted alkoxy  
 CC or amino, and A is an E2F family dimer forming region or DNA binding  
 CC region, of at least 12 consecutive amino acids. Compounds of this formula  
 CC can be used to inhibit E2F activity, and are useful in the treatment and  
 CC prevention of tumours and arteriosclerosis.

SX Sequence 28 AA;

Query Match 45.5%; Score 46; DB 19; Length 28;  
 Best Local Similarity 100.0%; Pred. No. 2.2;  
 Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 NVLMAMNIIS 10  
 |||||  
 Db 19 NVLMAMNIIS 28



## RESULT 17

AAW30517  
ID AAW30517 standard; Peptide; 19 AA.

XX AC AAW30517;  
XX DT 26-OCT-1998 (first entry)  
XX DE DP-1 transcription factor peptide H2mt3.

XX KW DP-1; transcription factor; antagonist; E2F protein; apoptosis;  
KW cell proliferation; cardiovascular cell; restenosis; tumour;  
KW surgical stent; therapy.

XX OS Synthetic.  
XX OS Homo sapiens.

XX FH Key Location/Qualifiers  
FT Misc-difference 8  
FT FT /note= "L173R mutation"  
FT FT Misc-difference 11  
FT FT /note= "L176R mutation"

XX PN WO9828334-A1.

XX PD 02-JUL-1998.

XX PF 22-DEC-1997; 97WO-GB03506.

XX PR 20-DEC-1996; 96GB-0026589.

XX PA (PROL-) PROLIFIX LTD.

XX PI Bandara LR, La Thangue NB;  
XX DR WPI; 1998-377596/32.

XX PT Polypeptide fragments of the DP-1 transcription factor - used for  
PT inducing apoptosis, specifically in tumour and cardiovascular cells,  
PT e.g. for preventing restenosis

XX PS Example D; Page 26; 55pp; English.

XX CC Peptide H2mt3 is based on peptide H2 (see AAW30504) from the DEF box  
CC (see AAW30501) of transcription factor DPl. In H2mt3, amino acid  
CC residues of H2 that correspond to DPl residues Leu173 and Leu176  
CC are substituted by Arg residues. H2 is an antagonist of the  
CC heterodimerisation of DPl with E2F. H2mt3 has none of the  
CC antagonistic activity of H2. H2 and other claimed peptides (see  
CC AAW30504-07) from the DEF box region of DPl can be used to induce  
CC apoptosis, specifically in tumour and cardiovascular cells, e.g.  
CC for the prevention of restenosis.

XX SQ Sequence 19 AA;

Query Match 44.6%; Score 45; DB 19; Length 19;

Best Local Similarity 90.9%; Pred. No. 2;

Matches 10; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 NVLMAMNIISK 11

DB 9 NVRMAMNIISK 19

## RESULT 18

AAU78095  
ID AAU78095 standard; Protein; 85 AA.

XX AC AAU78095;

XX DT 18-JUN-2002 (first entry)

XX Human DNA binding domain E2F-1.

XX Human; telomerase reverse transcriptase; TERT; Site C; Progeria; burn;  
KW repressor binding site; Hutchinson-Gilford syndrome; AIDS; cancer;  
KW acquired immunodeficiency syndrome; cardiovascular disease; osteoporosis;  
KW skin rejuvenation; immune senescence; bone marrow transplant; skin graft;  
KW neoplastic disease; TERT minimal promoter; DNA binding domain; E2F-1.

XX OS Homo sapiens.

XX FH Key Location/Qualifiers  
FT Domain 18..59  
FT FT /label= DNA\_binding\_domain

XX PN WO200216657-A1.

XX PD 28-FEB-2002.

XX PF 17-AUG-2001; 2001WO-US25861.

XX PR 24-AUG-2000; 2000US-227865P.

XX PR 01-SEP-2000; 2000US-230174P.

XX PR 05-OCT-2000; 2000US-238345P.

XX PA (SIER-) SIERRA SCI INC.

XX PI Andrews WH, Foster CA, Fraser S, Mohammadpour H;

XX DR WPI; 2002-280952/32.

XX PT Modulating expression of telomerase reverse transcriptase (TERT) in a  
PT cell, for regulating proliferative capacity of a cell, involves  
PT modulating TERT transcription repression by Site C repressor binding  
PT site -

XX PS Disclosure; Page 8; 66pp; English.

XX CC The present invention relates to a new method of modulating expression  
CC of telomerase reverse transcriptase (TERT) from a TERT expression system  
CC that includes a TERT promoter and a Site C repressor binding site. The  
CC method of the invention involves modulating TERT transcription repression  
CC by the Site C repressor binding site. The method of the invention is  
CC useful for modulating expression of TERT for producing a mammalian  
CC antibody. The method is also useful in a variety of different  
CC applications, including immortalisation of cells, production of reagents  
CC for use in life science research, therapeutic applications, and  
CC therapeutic agent screening applications. Increasing TERT expression  
CC delays natural telomeric shortening and/or increases telomeric length and  
CC is useful for treating disease conditions such as Progeria or  
CC Hutchinson-Gilford syndrome, acquired immunodeficiency syndrome (AIDS),  
CC cardiovascular disease, osteoporosis, in skin rejuvenation and to inhibit  
CC immune senescence. The method can be employed to lengthen telomeres of  
CC osteoblast and osteoclast stem cells, encouraging bone replacement and  
CC proper remodeling and reinforcement, and can thus be used in bone marrow  
CC transplants for the treatment of cancer and skin grafts for burn  
CC victims and as such the method improves the survival and effectiveness of  
CC bone marrow and skin cell transplants. Decreasing TERT expression is  
CC useful for treating cellular proliferative disease conditions, including  
CC neoplastic disease conditions e.g. cancer. The present amino acid  
CC sequence represents the human DNA binding domain E2F-1.

XX SQ Sequence 85 AA;

Query Match 44.1%; Score 44.5; DB 23; Length 85;

Best Local Similarity 42.9%; Pred. No. 13;

Matches 9; Conservative 6; Mismatches 5; Indels 1; Gaps 1;

QY 1 NVLMAMNIISK-KKEIKWIG 20

DB 63 NVLEGQLIAKSKNHIQWLG 83

RESULT 19  
 ABB82986  
 ID ABB82986 standard; Protein; 85 AA.  
 XX  
 AC ABB82986;  
 XX  
 DT 14-APR-2003 (first entry)  
 XX  
 DE Repressor protein E2F-1 DNA binding domain.  
 XX  
 KW Telomerase reverse transcriptase; TERT; Site C repressor; transcription;  
 KW cytotaxtic; immunostimulant; anti-HIV; vulnerary; telomerase; human;  
 KW repressor protein; E2F-1.  
 XX  
 OS Homo sapiens.  
 XX  
 PN WC2002101010-A2.  
 XX  
 PD 19-DEC-2002.  
 XX  
 PF 06-JUN-2002; 2002WO-US17959.  
 XX  
 PR 07-JUN-2001; 2001US-296992P.  
 XX  
 PA (SIER-) SIERRA SCI INC.  
 XX  
 PI Foster CA, Fraser S, Mohammadpour H, Andrews WH;  
 XX  
 DR WPI; 2003-167401/16.  
 XX  
 PT Modulating expression of telomerase reverse transcriptase TERT by  
 PT blocking expression of TERT transcription, useful for the diagnosis and  
 PT treatment of disorders associated with aberrant telomerase activity  
 PT such as cancer and HIV  
 XX  
 PS Disclosure; Page 7; 47pp; English.  
 XX  
 CC The invention relates to modulating expression of telomerase reverse  
 CC transcriptase (TERT) expression system that includes a TERT promoter and  
 CC a Site C repressor binding site. The method involves modulating TERT  
 CC transcription repression by the Site C repressor binding site. The  
 CC methods and compositions of the present invention are useful for the  
 CC immortalization of cells, production of reagents in life science  
 CC research, therapeutic agent screening applications, diagnosis and  
 CC treatment of disorders associated with aberrant telomerase activity such  
 CC as cancer, progeria, immune senescence, HIV, and in skin rejuvenation.  
 CC The present sequence represents the DNA binding domain of a repressor  
 CC protein E2F-1.  
 XX  
 SQ Sequence 85 AA;  
 Query Match 44.1%; Score 44.5; DB 24; Length 85;  
 Best Local Similarity 42.9%; Pred. No. 13;  
 Matches 9; Conservative 6; Mismatches 5; Indels 1; Gaps 1;  
 QY 1 NVLMANNISKE-KKEIKWIG 20  
 DB 63 NVLEGIQIJAKKSKNHIQWLG 83  
 RESULT 20  
 AAW30502  
 ID AAW30502 standard; Peptide; 9 AA.  
 XX  
 AC AAW30502;  
 XX  
 DT 26-OCT-1998 (first entry)  
 XX  
 DE DP-1 transcription factor antagonist peptide motif.  
 XX  
 KW DP-1; transcription factor; antagonist; E2F protein; apoptosis;  
 KW cell proliferation; cardiovascular cell; restenosis; tumour;  
 KW surgical stent; therapy.

XX Synthetic.  
 OS Homo sapiens.  
 XX  
 PN WO9828334-A1.  
 XX  
 PD 02-JUL-1998.  
 XX  
 PF 22-DEC-1997; 97WO-GB03506.  
 XX  
 PR 20-DEC-1996; 96GB-0026589.  
 XX  
 PA (PROL-) PROLIFIX LTD.  
 XX  
 PI Bandara LR, La Thangue NB;  
 XX  
 DR WPI; 1998-377596/32.  
 XX  
 PT Polypeptide fragments of the DP-1 transcription factor - used for  
 PT inducing apoptosis, specifically in tumour and cardiovascular cells,  
 PT e.g. for preventing restenosis  
 XX  
 PS Claim 3; Page 44; 55pp; English.  
 XX  
 CC This peptide comprises amino acid residues 175-183 in the DEF box  
 CC region (see AAW30501) of transcription factor DP1. Claimed peptides  
 CC (II) (see AAW30502-07) containing this and/or another motif (see  
 CC AAW30503) of the DEF box, are antagonists of the heterodimerisation  
 CC of a DP protein with an E2F protein. Also claimed are variants  
 CC of these peptides, especially containing substitutions of residues  
 CC corresponding to residues 167, 169, 171 and 175 of DP-1, fusion  
 CC proteins (III) comprising (I) or (II) and a membrane translocation  
 CC sequence (see AAW30508), expression vectors encoding (I)-(III) and  
 CC host cells. (I)-(III) are used therapeutically to induce apoptosis,  
 CC specifically in tumour or cardiovascular cells, either in vivo or in  
 CC vitro, e.g. for purging bone marrow. Surgical stents comprising  
 CC (I)-(III) are used to treat or prevent restenosis in patients who  
 CC have undergone angioplasty. (I)-(III) function by inactivating  
 CC the DNA-binding activity of DP/E2F heterodimers. They are also  
 CC used as research reagents, as positive controls in assays for  
 CC identifying antagonists of DP-1/E2F dimerisation and as immunoassay  
 CC agents. Also described is the use of sequences antisense to  
 CC nucleic acids encoding (I)-(III) to control DP levels in cells,  
 CC particularly by gene therapy. When formulated with cytotoxic  
 CC or cytostatic agents, (I)-(III) enhance cell killing.  
 XX  
 SQ Sequence 9 AA;  
 Query Match 41.6%; Score 42; DB 19; Length 9;  
 Best Local Similarity 100.0%; Pred. No. 9.3e+05;  
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 NVLMANNII 9  
 DB 1 NVLMANNII 9  
 RESULT 21  
 AAY37764  
 ID AAY37764 standard; Protein; 91 AA.  
 XX  
 AC AAY37764;  
 XX  
 DT 07-OCT-1999 (first entry)  
 XX  
 DE Amino acid sequence of a Chlamydia trachomatis protein.  
 XX  
 KW Vaccine; eye disease; conventional trachoma; nonendemic trachoma;  
 KW paratrachoma; inclusion conjunctivitis; genital disease; perihepatitis;  
 KW nongonococcal urethritis; epididymitis; cervicitis; salpingitis;  
 KW bartholinitis; pneumopathy; venereal lymphogranulomatosis.  
 XX  
 OS Chlamydia trachomatis.

[illegible]

neuroprotective; antiatherosclerotic; anticoagulant; thrombolytic;  
cardiant; hypotensive; antithyroid; antiinflammatory; immunomodulator;  
dermatological; analgesic; virucide; antibacterial; fungicide.  
Homo sapiens.  
WO200190366-A2.  
29-NOV-2001.  
24-MAY-2001; 2001WO-US17076.  
24-MAY-2000; 2000US-206690P.  
(CURA-) CURAGEN CORP.  
Leach MD, Shimkets RA;  
WPI; 2002-106200/14.  
N-PSDB; ABN76738.  
Novel human polypeptides and polynucleotides useful for diagnosing,  
preventing and treating cardiovascular disease, neurodegenerative,  
hyperproliferative disorders and disorders related to organ  
transplantation  
Claim 10; Page 1101; 2508pp; English.  
Sequences ABP31028-ABP3561 represent 4534 novel human proteins  
designated ORF (open reading frame) 1-4534, and sequences ABN75054-  
ABN79587 represent cDNAs encoding them. The invention also encompasses  
polypeptides at least 80% identical to the ORF1-ORF4534 (collectively to  
referred to as ORFX) proteins, polynucleotides at least 85% identical to  
the ORFX nucleic acid sequences, vectors and host cells comprising ORFX  
polynucleotides, the recombinant production of ORFX proteins, antibodies  
specific for ORFX proteins, methods of detecting ORFX polynucleotides and  
polypeptides, methods of screening for modulators of ORFX expression or  
activity, and methods of screening individuals for a predisposition to an  
ORFX-associated disorder. The ORFX proteins of the invention have a wide  
range of biological activities, such as cytokine, cell proliferation,  
cell differentiation, immune modulation, haematopoiesis regulation,  
tissue growth, angiogenesis, activin or inhibin activity, chemotactic/  
chemokinetic activity, haemostatic activity, thrombolytic activity,  
receptor/ligand, antiinflammatory activity, tumour inhibition activity,  
and antiinfective activity, and may also be involved in the determination  
of bodily characteristics, fertility and behaviour. ORFX proteins,  
nucleic acids and antibodies may be used in the treatment of cancers,  
other proliferative disorders such as epilepsy and Alzheimer's disease,  
neurological disorders such as epilepsy and Alzheimer's disease,  
cardiovascular diseases, immune system disorders, disorders related to  
organ transplantation, disorders of tissue growth and regeneration,  
diseases such as diabetes mellitus, hypothyroidism, and cholesterol ester  
storage disease, and infectious diseases caused by viral, bacterial,  
fungal and other pathogens. ORFX nucleic acids may also be used as a  
source of primers and probes, in the detection of ORFX genomic sequences  
or transcripts, in the identification and cloning of homologous  
sequences, in genetic diagnosis, and in forensic biology. The ORFX  
nucleic acids may additionally be used to produce transgenic animals  
which may be useful for studying the function and/or activity of ORFX  
protein, and in drug screening. The ORFX proteins may also be used as  
immunogens to generate specific antibodies, which are useful in the  
diagnosis, treatment and monitoring of ORFX-associated diseases.

Query Match 39.6%; Score 40; DB 23; Length 63;  
Best Local Similarity 50.0%; Pred. No. 50;  
Matches 6; Conservative 3; Mismatches 3; Indels 0; Gaps 0;  
QY 9 ISKEKKEIKWIG 20  
: ||| : : :  
48 VQESKENQWLG 59  
DB

RESULT 24  
ABG49671  
ID ABG49671 standard; Peptide; 75 AA.  
XX  
AC ABG49671;  
XX  
DT 25-FEB-2003 (first entry)  
XX  
DE Human liver peptide, SEQ ID NO 28319.  
XX  
KW Human; liver; cirrhosis; hyperlipoproteinaemia; hyperlipidaemia;  
hypercholesterolaemia; coronary heart disease.  
XX  
OS Homo sapiens.  
XX  
PN WO200157273-A2.  
XX  
PD 09-AUG-2001.  
XX  
PF 30-JAN-2001; 2001WO-US00664.  
XX  
PR 04-FEB-2000; 2000US-0180312.  
PR 26-MAY-2000; 2000US-0207456.  
PR 30-JUN-2000; 2000US-0608408.  
PR 03-AUG-2000; 2000US-0632366.  
PR 21-SEP-2000; 2000US-0234687.  
PR 27-SEP-2000; 2000US-0236359.  
PR 04-OCT-2000; 2000GB-0024263.  
XX  
(MOLE-) MOLECULAR DYNAMICS INC.  
XX  
PI Penn SG, Hanzel DK, Chen W, Rank DR;  
XX  
XX WPI; 2001-488898/53.  
XX  
PT Human genome-derived single exon nucleic acid probes useful for  
analysing gene expression in human adult liver -  
XX  
XX Claim 27; SEQ ID NO 28319; 658pp; English.  
XX  
XX The invention relates to a single exon nucleic acid probe (SENP) (I) for  
measuring human gene expression in a sample derived from human adult  
liver, comprising one of 13109 defined nucleotide sequences given in the  
specification (or complements/ fragments). The probe hybridises at high  
stringency to a nucleic acid molecule expressed in the human adult  
liver. (I) may be used for predicting, measuring and displaying gene  
expression in samples derived from human adult liver. The genes  
identified may be involved in genetic liver diseases such as cirrhosis,  
hyperlipoproteinaemia, hyperlipidaemia and hypercholesterolaemia which  
is associated with coronary heart disease. ABG47348-ABG59930 represent  
human liver single exon encoded peptides of the invention.  
XX Note: The sequence information for this patent does not appear in the  
printed specification but was obtained in electronic format directly  
from WIPO at ffp.wipo.int/pub/published\_pct\_sequences.  
XX  
SQ Sequence 75 AA;  
Query Match 39.6%; Score 40; DB 22; Length 75;  
Best Local Similarity 50.0%; Pred. No. 61;  
Matches 6; Conservative 3; Mismatches 3; Indels 0; Gaps 0;  
QY 9 ISKEKKEIKWIG 20  
: ||| : : :  
31 ITKEKSLRWAG 42  
DB  
RESULT 25  
ABB29659  
ID ABB29659 standard; Peptide; 75 AA.  
XX  
AC ABB29659;  
XX

DT 01-FEB-2002 (first entry)  
XX Peptide #2310 encoded by breast cell single exon nucleic acid probe.  
DE Human; microarray; single exon probe; gene expression; breast;  
KW disease; cancer.  
KW Homo sapiens.  
OS WO200157271-A2.  
XX 09-AUG-2001.  
XX 30-JAN-2001; 2001WO-US00662.  
XX 04-FEB-2000; 2000US-0180312.  
XX 26-MAY-2000; 2000US-0207456.  
XX 30-JUN-2000; 2000US-0608408.  
XX 03-AUG-2000; 2000US-0632366.  
XX 21-SEP-2000; 2000US-0234687.  
XX 27-SEP-2000; 2000US-0236359.  
XX 04-OCT-2000; 2000GB-0024263.  
XX (MOLE-) MOLECULAR DYNAMICS INC.  
PA Penn SG, Hanzel DK, Chen W, Rank DR;  
XX WPI; 2001-496933/54.  
XX New spatially-addressable set of single exon nucleic acid probes,  
DR useful for measuring gene expression in sample derived from human  
XX breast, comprises number of single exon nucleic acid probes -  
XX Claim 27; SEQ ID NO 12627; 327pp + sequence listing; English.  
XX The invention relates to a spatially-addressable set of single exon  
CC nucleic acid probes for measuring gene expression in a sample derived  
CC from human breast and Bt 474 cells. The method involves contacting  
CC the probes with a collection of detectably labelled nucleic acids  
CC derived from mRNA of human breast, and then measuring the label  
CC bound to each probe of the microarray. The probes are useful for  
CC verifying the expression of regions of genomic DNA predicted to  
CC encode proteins. They are useful for gene discovery, and for  
CC determining predisposition and/or prognosing breast disease. Gene  
CC expression analysis is useful for assessing the toxicity of chemical  
CC agents on cells. The microarray of this invention presents a far greater  
CC diversity of probes for measuring gene expression, with far less bias  
CC than expressed sequence tag microarrays. The method is suitable for  
CC rapid production of functional information from genomic sequence. The  
CC present sequence is a peptide encoded by a single exon nucleic acid  
CC probe of the invention.  
CC Note: The sequence data for this patent did not form part of the  
CC printed specification, but was obtained in electronic format directly  
CC from WIPO at ftp.wipo.int/pub/published\_pct\_sequences.  
XX  
SQ Sequence 75 AA;  
Query Match 39.6%; Score 40; DB 22; Length 75;  
Best Local Similarity 50.0%; Pred. No. 61;  
Matches 6; Conservative 3; Mismatches 3; Indels 0; Gaps 0;  
QY 9 ISKEKKEIKWIG 20  
Db :||| :|||  
31 ITKEKSLRWAG 42  
RESULT 26  
ABB34837  
ID ABB34837 standard; Peptide; 75 AA.  
XX  
XX ABB34837;  
AC  
XX 04-FEB-2002 (first entry)  
DT

XX Peptide #2343 encoded by human foetal liver single exon probe.  
DE Human; foetal liver; gene expression; single exon nucleic acid probe.  
KW Homo sapiens.  
OS WO200157277-A2.  
XX 09-AUG-2001.  
XX 30-JAN-2001; 2001WO-US00669.  
XX 04-FEB-2000; 2000US-0180312.  
XX 26-MAY-2000; 2000US-0207456.  
XX 30-JUN-2000; 2000US-0608408.  
XX 03-AUG-2000; 2000US-0632366.  
XX 21-SEP-2000; 2000US-0234687.  
XX 27-SEP-2000; 2000US-0236359.  
XX 04-OCT-2000; 2000GB-0024263.  
XX (MOLE-) MOLECULAR DYNAMICS INC.  
PA Penn SG, Hanzel DK, Chen W, Rank DR;  
XX WPI; 2001-483447/52.  
XX Human genome-derived single exon nucleic acid probes useful for  
DR analyzing gene expression in human fetal liver -  
XX Claim 27; SEQ ID NO 27472; 639pp + sequence listing; English.  
XX The invention relates to a single exon nucleic acid probe for  
CC measuring human gene expression in a sample derived from human foetal  
CC liver. The single exon nucleic acid probes may be used for predicting,  
CC measuring and displaying gene expression in samples derived from human  
CC fetal liver. The present sequence is a peptide encoded by a single exon  
CC nucleic acid probe of the invention.  
CC Note: The sequence data for this patent did not form part of the  
CC printed specification, but was obtained in electronic format directly  
CC from WIPO at ftp.wipo.int/pub/published\_pct\_sequences.  
XX  
SQ Sequence 75 AA;  
Query Match 39.6%; Score 40; DB 22; Length 75;  
Best Local Similarity 50.0%; Pred. No. 61;  
Matches 6; Conservative 3; Mismatches 3; Indels 0; Gaps 0;  
QY 9 ISKEKKEIKWIG 20  
Db :||| :|||  
31 ITKEKSLRWAG 42  
RESULT 27  
ABB20257  
ID ABB20257 standard; Protein; 75 AA.  
XX  
XX ABB20257;  
AC  
XX 23-JAN-2002 (first entry)  
DT  
XX Protein #2256 encoded by probe for measuring heart cell gene expression.  
DE Human; gene expression; heart; microarray; vascular system;  
XX cardiovascular disease; hypertension; cardiac arrhythmia;  
KW congenital heart disease.  
XX Homo sapiens.  
OS WO200157274-A2.  
XX  
XX 09-AUG-2001.  
XX

PF 30-JAN-2001; 2001WO-US00666.  
XX  
PR 04-FEB-2000; 2000US-0180312.  
PR 26-MAY-2000; 2000US-0207456.  
PR 30-JUN-2000; 2000US-0608408.  
PR 03-AUG-2000; 2000US-0632366.  
PR 21-SEP-2000; 2000US-0234687.  
PR 27-SEP-2000; 2000US-0236359.  
PR 04-OCT-2000; 2000GB-0024263.  
XX  
PA (MOLE-) MOLECULAR DYNAMICS INC.  
XX  
PI Penn SG, Hanzel DK, Chen W, Rank DR;  
XX  
DR WPI; 2001-488899/53.  
XX  
PT Single exon nucleic acid probes for analyzing gene expression in human  
PT hearts -  
XX  
PS Claim 15; SEQ ID No 22027; 530pp; English.  
XX  
XX The present invention relates to single exon nucleic acid probes for  
CC measuring human gene expression in a sample derived from human heart (see  
CC ABA21535-ABA41305). The present sequence is a protein encoded by one such  
CC probe. The probes may be used for predicting, measuring and displaying  
CC gene expression in samples derived from the human heart via microarrays.  
CC By measuring gene expression, the probes are useful for predicting  
CC diagnosing, grading, staging, monitoring and prognosing diseases of the  
CC human heart and vascular system e.g. cardiovascular disease,  
CC hypertension, cardiac arrhythmias and congenital heart disease.  
CC Note: The sequence data for this patent did not form part of the printed  
CC specification, but was obtained in electronic format directly from WIPO  
CC at ftp.wipo.int/pub/published\_pct\_sequences.  
XX  
SQ Sequence 75 AA;  
Query Match 39.6%; Score 40; DB 22; Length 75;  
Best Local Similarity 50.0%; Pred. No. 61;  
Matches 6; Conservative 3; Mismatches 3; Indels 0; Gaps 0;  
Qy 9 ISKEKKEIKWIG 20  
Db 31 ITKEKSSLRWAG 42  
RESULT 28  
AAM55645  
ID AAM55645 standard; Protein; 75 AA.  
XX  
AC AAM55645;  
XX  
DT 05-NOV-2001 (first entry)  
XX  
DE Human brain expressed single exon probe encoded protein SEQ ID NO: 27750.  
XX  
KW Human; brain expressed exon; gene expression analysis; probe;  
KW microarray; Alzheimer's disease; multiple sclerosis; schizophrenia;  
KW epilepsy; cancer.  
XX  
OS Homo sapiens.  
XX  
PN WO200157275-A2.  
XX  
PD 09-AUG-2001.  
XX  
PF 30-JAN-2001; 2001WO-US00667.  
XX  
PR 04-FEB-2000; 2000US-0180312.  
PR 26-MAY-2000; 2000US-0207456.  
PR 30-JUN-2000; 2000US-0608408.  
PR 03-AUG-2000; 2000US-0632366.  
PR 21-SEP-2000; 2000US-0234687.  
PR 27-SEP-2000; 2000US-0236359.  
XX

PR 04-OCT-2000; 2000GB-0024263.  
XX  
PA (MOLE-) MOLECULAR DYNAMICS INC.  
XX  
PI Penn SG, Hanzel DK, Chen W, Rank DR;  
XX  
DR WPI; 2001-483446/52.  
XX  
PT Single exon nucleic acid probes for analyzing gene expression in human  
PT brains -  
XX  
PS Example 4; SEQ ID NO: 27750; 650pp + Sequence Listing; English.  
XX  
XX The present invention provides a number of single exon nucleic acid  
CC probes which are derived from genomic sequences expressed in the human  
CC brain. They can be used to measure gene expression in brain cell samples,  
CC which may enable the diagnosis and improved treatment of nervous system  
CC diseases such as Alzheimer's disease, multiple sclerosis, schizophrenia, of  
CC epilepsy and cancers. The present sequence is a protein encoded by one of  
CC the probes of the invention.  
XX  
SQ Sequence 75 AA;  
Query Match 39.6%; Score 40; DB 22; Length 75;  
Best Local Similarity 50.0%; Pred. No. 61;  
Matches 6; Conservative 3; Mismatches 3; Indels 0; Gaps 0;  
Qy 9 ISKEKKEIKWIG 20  
Db 31 ITKEKSSLRWAG 42  
RESULT 29  
AAM68025  
ID AAM68025 standard; Protein; 75 AA.  
XX  
AC AAM68025;  
XX  
DT 06-NOV-2001 (first entry)  
XX  
DE Human bone marrow expressed probe encoded protein SEQ ID NO: 28331.  
XX  
KW Human; bone marrow expressed exon; gene expression analysis; probe;  
KW microarray; cancer; leukaemia; lymphoma; myeloma.  
XX  
OS Homo sapiens.  
XX  
PN WO200157276-A2.  
XX  
PD 09-AUG-2001.  
XX  
PF 30-JAN-2001; 2001WO-US00668.  
XX  
PR 04-FEB-2000; 2000US-0180312.  
PR 26-MAY-2000; 2000US-0207456.  
PR 30-JUN-2000; 2000US-0608408.  
PR 03-AUG-2000; 2000US-0632366.  
PR 21-SEP-2000; 2000US-0234687.  
PR 27-SEP-2000; 2000US-0236359.  
PR 04-OCT-2000; 2000GB-0024263.  
XX  
PA (MOLE-) MOLECULAR DYNAMICS INC.  
XX  
PI Penn SG, Hanzel DK, Chen W, Rank DR;  
XX  
DR WPI; 2001-488900/53.  
XX  
PT Human genome-derived single exon nucleic acid probes useful for  
PT analyzing gene expression in human bone marrow -  
XX  
PS Example 4; SEQ ID NO: 28331; 658pp + Sequence Listing; English.  
XX  
XX The present invention provides a number of single exon nucleic acid  
CC

CC probes which are derived from genomic sequences expressed in the human  
 CC bone marrow. They can be used to measure gene expression in bone marrow  
 CC samples, which may enable the improved diagnosis and treatment of cancers  
 CC such as lymphoma, leukaemia and myeloma. The present sequence is a  
 CC protein encoded by one of the probes of the invention.

XX SQ Sequence 75 AA;

Query Match 39.6%; Score 40; DB 22; Length 75;

Best Local Similarity 50.0%; Pred. No. 61;

Matches 6; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

Qy 9 ISKEKKEIKWIG 20

Db 31 ITKEKSSLRWAG 42

RESULT 30

AA015843  
 ID AAM15843 standard; Protein; 75 AA.

XX AC AAM15843;

XX DT 12-OCT-2001 (first entry)

XX DE Peptide #2277 encoded by probe for measuring cervical gene expression.

XX KW Probe; human; microarray; gene expression; cervical epithelial cell;  
 cervical cancer.

XX OS Homo sapiens.

XX PN WO200157278-A2.

XX PD 09-AUG-2001.

XX PF 30-JAN-2001; 2001WO-US000670.

XX PR 04-FEB-2000; 2000US-0180312.

XX PR 26-MAY-2000; 2000US-0207456.

XX PR 30-JUN-2000; 2000US-0608408.

XX PR 03-AUG-2000; 2000US-0632366.

XX PR 21-SEP-2000; 2000US-0234687.

XX PR 27-SEP-2000; 2000US-0236359.

XX PR 04-OCT-2000; 2000GB-0024263.

XX PA (MOLE-) MOLECULAR DYNAMICS INC.

XX PI Penn SG, Hanzel DK, Chen W, Rank DR;

XX DR WPI; 2001-488901/53.

XX PT Human genome-derived single exon nucleic acid probes useful for  
 PT analyzing gene expression in human cervical epithelial cells -

XX PS Claim 27; SEQ ID No 20669; 487pp; English.

XX CC The present invention relates to human single exon nucleic acid probes  
 CC (SENPs; see AAL10068-AAL28459). The present sequence is a peptide encoded  
 CC by one such probe. The SENPs are derived from human Hela cells. The SENPs  
 CC can be used to produce a single exon microarray, which can be used for  
 CC measuring human gene expression in a sample derived from human cervical  
 CC epithelial cells. By measuring gene expression, the probes are therefore  
 CC useful in grading and/or staging of diseases of the cervix, notably  
 CC cervical cancer.

CC Note: The sequence data for this patent did not form part of the printed  
 CC specification, but was obtained in electronic format directly from WIPO  
 CC at ftp.wipo.int/pub/published\_pct\_sequences.

XX SQ Sequence 75 AA;

Query Match 39.6%; Score 40; DB 22; Length 75;

Best Local Similarity 50.0%; Pred. No. 61;

Matches 6; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

Qy 9 ISKEKKEIKWIG 20

Db 31 ITKEKSSLRWAG 42

RESULT 31

AA028353  
 ID AAM28353 standard; Protein; 75 AA.

XX AC AAM28353;

XX DT 17-OCT-2001 (first entry)

XX DE Peptide #2390 encoded by probe for measuring placental gene expression.

XX KW Probe; microarray; human; placenta; antenatal diagnosis;  
 genetic disorder.

XX OS Homo sapiens.

XX PN WO200157272-A2.

XX PD 09-AUG-2001.

XX PF 30-JAN-2001; 2001WO-US000663.

XX PR 04-FEB-2000; 2000US-0180312.

XX PR 26-MAY-2000; 2000US-0207456.

XX PR 30-JUN-2000; 2000US-0608408.

XX PR 03-AUG-2000; 2000US-0632366.

XX PR 21-SEP-2000; 2000US-0234687.

XX PR 27-SEP-2000; 2000US-0236359.

XX PR 04-OCT-2000; 2000GB-0024263.

XX PA (MOLE-) MOLECULAR DYNAMICS INC.

XX PI Penn SG, Hanzel DK, Chen W, Rank DR;

XX DR WPI; 2001-488901/53.

XX PT Human genome-derived single exon nucleic acid probes useful for  
 PT analyzing gene expression in human placenta -

XX PS Claim 27; SEQ ID No 28622; 654pp; English.

XX CC The present invention relates to single exon nucleic acid probes (SENPs;  
 CC see AAL1315-AA157546). The present sequence is a peptide encoded by one  
 CC such probe. The probes are useful for producing a microarray for  
 CC predicting, measuring and displaying gene expression in samples derived  
 CC from human placenta. The probes are useful for antenatal diagnosis of  
 CC human genetic disorders.

XX SQ Sequence 75 AA;

Query Match 39.6%; Score 40; DB 22; Length 75;

Best Local Similarity 50.0%; Pred. No. 61;

Matches 6; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

Qy 9 ISKEKKEIKWIG 20

Db 31 ITKEKSSLRWAG 42

RESULT 32

AA03581  
 ID AAM03581 standard; Protein; 75 AA.

XX AC AAM03581;

XX DT 09-OCT-2001 (first entry)





Query Match 39.6%; Score 40; DB 23; Length 75;  
Best Local Similarity 50.0%; Pred. No. 61;  
Matches 6; Conservative 3; Mismatches 3; Indels

Qy 9 ISKEKEIKWIG 20  
|:|:|:|:|:|:|  
Db 31 ITKEKSSLRWAG 42

RESULT 34  
AAO01170  
ID AAO01170 standard; Protein; 83 AA.

XX	AAO01170;
XX	AC
XX	
DT	06-NOV-2001 (first entry)
XX	
DE	Human polypeptide SEQ ID NO 15062.

Human; cytokine; cell proliferation; cell differentiation; gene therapy;  
KW vaccine; peptide therapy; stem cell growth factor; haematopoiesis;  
KW tissue growth factor; immunomodulatory; cancer; leukaemia;  
KW nervous system disorders; arthritis; inflammation.

OS Homo sapiens.

AA  
PN  
WO200164835-A2.

07-SEP-2001.

PF 26-FEB-2001; 2001WO-US04927.

PR 28-FEB-2000; 2000US-0515126.

YR 10007-YM1-0T 10003-037409.

PA (HYSE-) HYSEQ INC.

PI Tang YT, Liu C, Drmanac RT;

DR WPI; 2001-514838/56.

10-10-68

PT Isolated nucleic acids and polypeptides, useful for preventing  
PT diagnosing and treating e.g. leukaemia, inflammation and immune  
PT disorders -

PS Claim 20; SEQ ID NO 15062; 1399pp + Sequence Listing: English.

The invention relates to human polynucleotides (AA179941-AA193841) and the encoded proteins (AAO00010-AAO13910) that exhibit activity elating to cytokine, cell proliferation or cell differentiation or which may induce production of other cytokines in other cell populations. The polynucleotides and polypeptides are useful in gene therapy, vaccines or peptide therapy. The polypeptides have various cytokine-like activities, e.g. stem cell growth factor activity, haematopoiesis regulating activity, tissue growth factor activity, immunomodulatory activity and activin/inhibin activity and may be useful in the diagnosis and/or treatment of cancer, leukaemia, nervous system disorders, arthritis and inflammation.

CC Note: The sequence data for this patent did not form part of the printed  
CC specification, but was obtained in electronic format directly from WIPO  
CC at [ftp.wipo.int/pub/published pct sequences](http://ftp.wipo.int/pub/published/pct/sequences).

Sequence 83 AA;

Query Match 39.6%; Score 40; DB 22; Length 83;  
Best Local Similarity 58.3%; Pred. No. 69;  
Matches 7; Conservative 4; Mismatches 1; Indels

Qy 9 ISKEKEIKWIG 20  
|||: ||: ||: ||

p6 9 ISKDLKELKYL 20

PR	08-SEP-2000;	2000US-0232081.
PR	12-SEP-2000;	2000US-0231968.
PR	14-SEP-2000;	2000US-0232397.
PR	14-SEP-2000;	2000US-0232398.
PR	14-SEP-2000;	2000US-0232399.
PR	14-SEP-2000;	2000US-0232400.
PR	14-SEP-2000;	2000US-0232401.
PR	14-SEP-2000;	2000US-0233063.
PR	14-SEP-2000;	2000US-0233064.
PR	14-SEP-2000;	2000US-0233065.
PR	21-SEP-2000;	2000US-0234223.
PR	21-SEP-2000;	2000US-0234274.
PR	23-SEP-2000;	2000US-0234397.
PR	23-SEP-2000;	2000US-0234998.
PR	26-SEP-2000;	2000US-0235484.
PR	27-SEP-2000;	2000US-0235834.
PR	27-SEP-2000;	2000US-0235836.
PR	29-SEP-2000;	2000US-0236327.
PR	29-SEP-2000;	2000US-0236367.
PR	29-SEP-2000;	2000US-0236368.
PR	29-SEP-2000;	2000US-0236369.
PR	29-SEP-2000;	2000US-0236370.
PR	02-OCT-2000;	2000US-0236802.
PR	02-OCT-2000;	2000US-0237037.
PR	02-OCT-2000;	2000US-0237038.
PR	02-OCT-2000;	2000US-0237039.
PR	02-OCT-2000;	2000US-0237040.
PR	13-OCT-2000;	2000US-0239935.
PR	20-OCT-2000;	2000US-0239937.
PR	20-OCT-2000;	2000US-0240960.
PR	20-OCT-2000;	2000US-0241221.
PR	20-OCT-2000;	2000US-0241785.
PR	20-OCT-2000;	2000US-0241786.
PR	20-OCT-2000;	2000US-0241787.
PR	20-OCT-2000;	2000US-0241808.
PR	20-OCT-2000;	2000US-0241809.
PR	20-OCT-2000;	2000US-0241826.
PR	01-NOV-2000;	2000US-024617.
PR	08-NOV-2000;	2000US-0246474.
PR	08-NOV-2000;	2000US-0246475.
PR	08-NOV-2000;	2000US-0246476.
PR	08-NOV-2000;	2000US-0246477.
PR	08-NOV-2000;	2000US-0246478.
PR	08-NOV-2000;	2000US-0246523.
PR	08-NOV-2000;	2000US-0246524.
PR	08-NOV-2000;	2000US-0246525.
PR	08-NOV-2000;	2000US-0246526.
PR	08-NOV-2000;	2000US-0246527.
PR	08-NOV-2000;	2000US-0246528.
PR	08-NOV-2000;	2000US-0246532.
PR	08-NOV-2000;	2000US-0246533.
PR	08-NOV-2000;	2000US-0246569.
PR	08-NOV-2000;	2000US-0246610.
PR	08-NOV-2000;	2000US-0246611.
PR	08-NOV-2000;	2000US-0246613.
PR	17-NOV-2000;	2000US-0249207.
PR	17-NOV-2000;	2000US-0249208.
PR	17-NOV-2000;	2000US-0249209.
PR	17-NOV-2000;	2000US-0249210.
PR	17-NOV-2000;	2000US-0249211.
PR	17-NOV-2000;	2000US-0249212.
PR	17-NOV-2000;	2000US-0249213.
PR	17-NOV-2000;	2000US-0249214.
PR	17-NOV-2000;	2000US-0249215.
PR	17-NOV-2000;	2000US-0249216.
PR	17-NOV-2000;	2000US-0249217.
PR	17-NOV-2000;	2000US-0249218.
PR	17-NOV-2000;	2000US-0249244.
PR	17-NOV-2000;	2000US-0249245.
PR	17-NOV-2000;	2000US-0249264.
PR	17-NOV-2000;	2000US-0249265.
PR	17-NOV-2000;	2000US-0249297.
PR	17-NOV-2000;	2000US-0249299.
PR	17-NOV-2000;	2000US-0249300.

PR	01-DEC-2000;	2000US-0250160.
PR	01-DEC-2000;	2000US-0250391.
PR	05-DEC-2000;	2000US-0251030.
PR	05-DEC-2000;	2000US-0251988.
PR	06-DEC-2000;	2000US-0256719.
PR	06-DEC-2000;	2000US-0251479.
PR	08-DEC-2000;	2000US-0251856.
PR	08-DEC-2000;	2000US-0251868.
PR	08-DEC-2000;	2000US-0251869.
PR	08-DEC-2000;	2000US-0251989.
PR	08-DEC-2000;	2000US-0251990.
PR	11-DEC-2000;	2000US-0254097.
PR	05-JAN-2001;	2001US-0259678.
XX	(HUMA-) HUMAN GENOME SCI INC.	
PI	Rosen CA, Barash SC, Ruben SM;	
DR	N-PSDB; AAK61062.	
XX	WPI; 2001-483426/52.	
XX	Nucleic acids encoding human immune/hematopoietic antigen polypeptides,	
PT	useful for preventing, diagnosing and/or treating cancers and	
PT	metastasis -	
XX	Claim 11; SEQ ID NO 15874; 3071pp + Sequence Listing; English.	
XX	AAK54951 to AAK64702 encode the human immune/haematopoietic antigen (I)	
CC	amino acid sequences given in AAM82170 to AAM91921. (I) have cytosstatic	
CC	activity, and can be used in gene therapy and vaccine production. (I)	
CC	proteins and polynucleotides may be used in the prevention, diagnosis and	
CC	treatment of diseases associated with inappropriate (I) expression. For	
CC	example, they may be used to treat disorders associated with decreased	
CC	expression by rectifying mutations or deletions in a patient's genome	
CC	that affect the activity of (I) by expressing inactive proteins or to	
CC	supplement the patients own production of (I). Additionally, (I)	
CC	polynucleotides may be used to produce the secreted (I), by inserting	
CC	the nucleic acids into a host cell and culturing the cell to express the	
CC	protein. (I) proteins and polynucleotides may be used to prevent,	
CC	diagnose and treat immune/haematopoietic-related diseases, especially	
CC	cancers and cancer metastases of haematopoietic-derived cells. AAK64703	
CC	to AAK87694 represent human immune/haematopoietic antigen genomic	
CC	sequences from the present invention. AAK54942 to AAK54950 and AAM82169	
CC	represent sequences used in the exemplification of the present invention.	
XX	Sequence 42 AA;	
SQ		
Query Match 39.1%; Score 39.5; DB 22; Length 42;		
Best Local Similarity 26.7%; Pred. No. 38;		
Matches 8; Conservative 8; Mismatches 3; Indels 11; Gaps 1;		
QY	1 NVLMMNIIISKE-----KKEIKWI 19	
DB	13 NFXISINLISETRNSLYITNALRKFKWL 42	
RESULT 36		
AAG09353		
ID	AAG09353 standard; Protein; 84 AA.	
XX	AC AAG09353;	
DT	17-OCT-2000 (first entry)	
XX	Arabidopsis thaliana protein fragment SEQ ID NO: 7253.	
XX	Protein identification; signal transduction pathway; metabolic pathway;	
KW	hybridisation assay; genetic mapping; gene expression control; promoter;	
KW	termination sequence.	
XX	Arabidopsis thaliana.	
OS		
XX	EPI033405-A2.	
PN		

XX 06-SEP-2000. 99US-0142055.  
PD 06-JUL-1999; 99US-0142390.  
XX 08-JUL-1999; 99US-0142803.  
PF 09-JUL-1999; 99US-0142920.  
XX 12-JUL-1999; 99US-0142977.  
PR 13-JUL-1999; 99US-0143542.  
PR 14-JUL-1999; 99US-0143624.  
PR 15-JUL-1999; 99US-0144005.  
PR 16-JUL-1999; 99US-0144085.  
PR 16-JUL-1999; 99US-0144086.  
PR 19-JUL-1999; 99US-0144325.  
PR 19-JUL-1999; 99US-0144331.  
PR 19-JUL-1999; 99US-0144332.  
PR 19-JUL-1999; 99US-0144333.  
PR 19-JUL-1999; 99US-0144334.  
PR 19-JUL-1999; 99US-0144335.  
PR 20-JUL-1999; 99US-0144352.  
PR 20-JUL-1999; 99US-0144632.  
PR 20-JUL-1999; 99US-0144884.  
PR 21-JUL-1999; 99US-0144814.  
PR 21-JUL-1999; 99US-0145086.  
PR 21-JUL-1999; 99US-0145088.  
PR 22-JUL-1999; 99US-0145085.  
PR 22-JUL-1999; 99US-0145087.  
PR 22-JUL-1999; 99US-0145089.  
PR 22-JUL-1999; 99US-0145192.  
PR 23-JUL-1999; 99US-0145145.  
PR 23-JUL-1999; 99US-0145218.  
PR 23-JUL-1999; 99US-0145224.  
PR 26-JUL-1999; 99US-0145276.  
PR 27-JUL-1999; 99US-0145913.  
PR 27-JUL-1999; 99US-0145918.  
PR 27-JUL-1999; 99US-0145919.  
PR 28-JUL-1999; 99US-0145951.  
PR 02-AUG-1999; 99US-0146386.  
PR 02-AUG-1999; 99US-0146388.  
PR 02-AUG-1999; 99US-0146389.  
PR 03-AUG-1999; 99US-0147038.  
PR 04-AUG-1999; 99US-0147204.  
PR 04-AUG-1999; 99US-0147302.  
PR 05-AUG-1999; 99US-0147192.  
PR 05-AUG-1999; 99US-0147260.  
PR 06-AUG-1999; 99US-0147303.  
PR 06-AUG-1999; 99US-0147416.  
PR 09-AUG-1999; 99US-0147493.  
PR 09-AUG-1999; 99US-0147935.  
PR 10-AUG-1999; 99US-0148171.  
PR 11-AUG-1999; 99US-0148319.  
PR 12-AUG-1999; 99US-0148341.  
PR 13-AUG-1999; 99US-0148565.  
PR 13-AUG-1999; 99US-0148684.  
PR 16-AUG-1999; 99US-0149368.  
PR 17-AUG-1999; 99US-0149175.  
PR 18-AUG-1999; 99US-0149426.  
PR 20-AUG-1999; 99US-0149722.  
PR 20-AUG-1999; 99US-0149723.  
PR 20-AUG-1999; 99US-0149929.  
PR 23-AUG-1999; 99US-0149902.  
PR 23-AUG-1999; 99US-0149930.  
PR 23-AUG-1999; 99US-0150566.  
PR 26-AUG-1999; 99US-0150884.  
PR 27-AUG-1999; 99US-0151065.  
PR 27-AUG-1999; 99US-0151066.  
PR 27-AUG-1999; 99US-0151080.  
PR 30-AUG-1999; 99US-0151303.  
PR 31-AUG-1999; 99US-0151438.  
PR 01-SEP-1999; 99US-0151930.  
PR 07-SEP-1999; 99US-0152363.  
PR 10-SEP-1999; 99US-0153070.  
PR 13-SEP-1999; 99US-0153758.  
PR 15-SEP-1999; 99US-0154018.  
PR 16-SEP-1999; 99US-0154039.  
PR 20-SEP-1999; 99US-0154779.

```
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159594.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161922.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

Query Match 38.6%; Score 39; DB 21; Length 84;
Best Local Similarity 42.1%; Pred. No. le+02;
Matches 8; Conservative 5; Mismatches 4; Indels 2; Gaps 1;

Qy 3 LMANNIISKEKKEIK-WI 19
Db 56 LIARRVVGREAKEIERVWI 74

RESULT 37
AM90074
ID AM90074 standard; Protein; 91 AA.
AC AM90074;
XX
XX
DT 07-NOV-2001 (first entry)
XX
DE Human immune/haematopoietic antigen SEQ ID NO:17667.
XX
KW Human; immune; haematopoietic; immune/haematopoietic antigen; cancer;
KW cytostatic; gene therapy; vaccine; metastasis.
XX
OS Homo sapiens.
XX
PN WO200157182-A2.
XX
PD 09-AUG-2001.
XX
PF 17-JAN-2001; 2001WO-US01354.
XX
XX 31-JAN-2000; 2000US-0179065.
PR 04-FEB-2000; 2000US-0180628.
PR 24-FEB-2000; 2000US-0184664.
PR 02-MAR-2000; 2000US-0186350.
PR 16-MAR-2000; 2000US-0189874.
PR 17-MAR-2000; 2000US-0190076.
PR 18-APR-2000; 2000US-0198123.
PR 19-MAY-2000; 2000US-0205515.
PR 07-JUN-2000; 2000US-0209467.
PR 28-JUN-2000; 2000US-0214886.
PR 30-JUN-2000; 2000US-0215135.
PR 07-JUL-2000; 2000US-0216647.
PR 07-JUL-2000; 2000US-0216880.
PR 11-JUL-2000; 2000US-0217487.
PR 11-JUL-2000; 2000US-0217496.
PR 14-JUL-2000; 2000US-0218290.
PR 26-JUL-2000; 2000US-0220963.
PR 26-JUL-2000; 2000US-0220964.
PR 14-AUG-2000; 2000US-0224518.
PR 14-AUG-2000; 2000US-0224519.
PR 14-AUG-2000; 2000US-0225213.
PR 14-AUG-2000; 2000US-0225214.
PR 14-AUG-2000; 2000US-0225266.
PR 14-AUG-2000; 2000US-0225267.
PR 14-AUG-2000; 2000US-0225268.
PR 14-AUG-2000; 2000US-0225270.
PR 14-AUG-2000; 2000US-0225447.
PR 14-AUG-2000; 2000US-0225757.
PR 14-AUG-2000; 2000US-0225758.
PR 14-AUG-2000; 2000US-0225759.
PR 18-AUG-2000; 2000US-0226279.
PR 22-AUG-2000; 2000US-0226681.
PR 22-AUG-2000; 2000US-0228868.
PR 22-AUG-2000; 2000US-0227182.
PR 23-AUG-2000; 2000US-0227009.
PR 23-AUG-2000; 2000US-0228924.
PR 01-SEP-2000; 2000US-0229287.
PR 01-SEP-2000; 2000US-0229343.
PR 01-SEP-2000; 2000US-0229344.
PR 01-SEP-2000; 2000US-0229345.
PR 05-SEP-2000; 2000US-0229509.
PR 05-SEP-2000; 2000US-0229513.
PR 06-SEP-2000; 2000US-0230437.
PR 06-SEP-2000; 2000US-0230438.
PR 08-SEP-2000; 2000US-0231242.
PR 08-SEP-2000; 2000US-0231243.
PR 08-SEP-2000; 2000US-0231244.
PR 08-SEP-2000; 2000US-0231413.
PR 08-SEP-2000; 2000US-0231414.
PR 08-SEP-2000; 2000US-0232080.
PR 08-SEP-2000; 2000US-0232081.
PR 12-SEP-2000; 2000US-0231968.
PR 14-SEP-2000; 2000US-0232397.
PR 14-SEP-2000; 2000US-0232398.
PR 14-SEP-2000; 2000US-0232399.
PR 14-SEP-2000; 2000US-0232400.
PR 14-SEP-2000; 2000US-0232401.
PR 14-SEP-2000; 2000US-0233063.
PR 14-SEP-2000; 2000US-0233064.
PR 14-SEP-2000; 2000US-0233065.
PR 21-SEP-2000; 2000US-0234223.
PR 21-SEP-2000; 2000US-0234274.
PR 25-SEP-2000; 2000US-0234997.
PR 25-SEP-2000; 2000US-0234998.
PR 25-SEP-2000; 2000US-0234998.
PR 26-SEP-2000; 2000US-0235484.
PR 27-SEP-2000; 2000US-0235834.
PR 27-SEP-2000; 2000US-0235836.
PR 29-SEP-2000; 2000US-0236327.
PR 29-SEP-2000; 2000US-0236367.
PR 29-SEP-2000; 2000US-0236368.
PR 29-SEP-2000; 2000US-0236369.
PR 29-SEP-2000; 2000US-0236370.
PR 02-OCT-2000; 2000US-0237037.
PR 02-OCT-2000; 2000US-0237038.
PR 02-OCT-2000; 2000US-0237039.
```

PR 02-OCT-2000; 2000US-0237040.  
PR 13-OCT-2000; 2000US-0239935.  
PR 13-OCT-2000; 2000US-0239937.  
PR 20-OCT-2000; 2000US-0240960.  
PR 20-OCT-2000; 2000US-0241221.  
PR 20-OCT-2000; 2000US-0241785.  
PR 20-OCT-2000; 2000US-0241786.  
PR 20-OCT-2000; 2000US-0241787.  
PR 20-OCT-2000; 2000US-0241808.  
PR 20-OCT-2000; 2000US-0241809.  
PR 01-NOV-2000; 2000US-0241826.  
PR 01-NOV-2000; 2000US-0244617.  
PR 08-NOV-2000; 2000US-0246474.  
PR 08-NOV-2000; 2000US-0246475.  
PR 08-NOV-2000; 2000US-0246476.  
PR 08-NOV-2000; 2000US-0246477.  
PR 08-NOV-2000; 2000US-0246478.  
PR 08-NOV-2000; 2000US-0246523.  
PR 08-NOV-2000; 2000US-0246524.  
PR 08-NOV-2000; 2000US-0246525.  
PR 08-NOV-2000; 2000US-0246526.  
PR 08-NOV-2000; 2000US-0246527.  
PR 08-NOV-2000; 2000US-0246528.  
PR 08-NOV-2000; 2000US-0246532.  
PR 08-NOV-2000; 2000US-0246532.  
PR 08-NOV-2000; 2000US-0246609.  
PR 08-NOV-2000; 2000US-0246610.  
PR 08-NOV-2000; 2000US-0246611.  
PR 08-NOV-2000; 2000US-0246613.  
PR 17-NOV-2000; 2000US-0249207.  
PR 17-NOV-2000; 2000US-0249208.  
PR 17-NOV-2000; 2000US-0249209.  
PR 17-NOV-2000; 2000US-0249210.  
PR 17-NOV-2000; 2000US-0249211.  
PR 17-NOV-2000; 2000US-0249212.  
PR 17-NOV-2000; 2000US-0249213.  
PR 17-NOV-2000; 2000US-0249214.  
PR 17-NOV-2000; 2000US-0249215.  
PR 17-NOV-2000; 2000US-0249216.  
PR 17-NOV-2000; 2000US-0249217.  
PR 17-NOV-2000; 2000US-0249218.  
PR 17-NOV-2000; 2000US-0249244.  
PR 17-NOV-2000; 2000US-0249245.  
PR 17-NOV-2000; 2000US-0249264.  
PR 17-NOV-2000; 2000US-0249265.  
PR 17-NOV-2000; 2000US-0249297.  
PR 17-NOV-2000; 2000US-0249299.  
PR 17-NOV-2000; 2000US-0249300.  
PR 01-DEC-2000; 2000US-0250160.  
PR 01-DEC-2000; 2000US-0250391.  
PR 05-DEC-2000; 2000US-0251030.  
PR 05-DEC-2000; 2000US-0251988.  
PR 05-DEC-2000; 2000US-0256719.  
PR 06-DEC-2000; 2000US-0251479.  
PR 08-DEC-2000; 2000US-0251856.  
PR 08-DEC-2000; 2000US-0251868.  
PR 08-DEC-2000; 2000US-0251869.  
PR 08-DEC-2000; 2000US-0251989.  
PR 08-DEC-2000; 2000US-0251990.  
PR 11-DEC-2000; 2000US-0254097.  
PR 03-JAN-2001; 2001US-02559678.  
XX  
PA (HUMA-) HUMAN GENOME SCI INC.  
XX  
XX Rosen CA, Barash SC, Ruben SM;  
PI  
XX  
DR WPI; 2001-483426/52.  
DR N-PSDB; AAKG2855.  
XX  
XX Nucleic acids encoding human immune/hematopoietic antigen polypeptides,  
PT useful for preventing, diagnosing and/or treating cancers and  
PT metastasis -  
XX  
XX Claim 11; SEQ ID NO 17667; 3071pp + Sequence Listing; English.

XX AAK54951 to AAK64702 encode the human immune/haematopoietic antigen (I)  
CC amino acid sequences given in AAM82170 to AAM91921. (I) have cytostatic  
CC activity, and can be used in gene therapy and vaccine production. (I)  
CC proteins and polynucleotides may be used in the prevention, diagnosis and  
CC treatment of diseases associated with inappropriate (I) expression. For  
CC example, they may be used to treat disorders associated with decreased  
CC expression by rectifying mutations or deletions in a patient's genome  
CC that affect the activity of (I) by expressing inactive proteins or to  
CC supplement the patients own production of (I). Additionally, (I)  
CC polynucleotides may be used to produce the secreted (I), by inserting  
CC the nucleic acids into a host cell and culturing the cell to express the  
CC protein. (I) proteins and polynucleotides may be used to prevent,  
CC diagnose and treat immune/haematopoietic-related diseases, especially  
CC cancers and cancer metastases of haematopoietic-derived cells. AAK64703  
CC to AAK87694 represent human immune/haematopoietic antigen genomic  
CC sequences from the present invention. AAK54942 to AAK54950 and AAM82169  
CC represent sequences used in the exemplification of the present invention.  
XX  
SQ Sequence 91 AA;  
Query Match 38.6%; Score 39; DB 22; Length 91;  
Best Local Similarity 66.7%; Pred. No. 1.1e+02;  
Matches 6; Conservative 3; Mismatches 0; Indels 0; Gaps 0;  
QY 10 SKEKEIKW 18  
Db 7 NKEKEQIKW 15  
:||||:|  
:  
RESULT 38  
ABP69141  
ID ABP69141 standard; Protein; 97 AA.  
XX  
XX AC ABP69141;  
XX  
DT 20-JAN-2003 (first entry)  
DE Human polypeptide SEQ ID NO 1188.  
XX  
XX Human; genome mapping; gene therapy; food supplement; virus; fungus;  
KW cell-proliferative disorder; neurodegenerative disease; bacterial;  
KW Parkinson's disease; Alzheimer's disease; autoimmune disease;  
KW Multiple sclerosis; diabetes; genetic disorder; wound; burn; infection;  
KW Arthritis; cytostatic; immunomodulator; nontropic; neuroprotective;  
KW antiparkinsonian; antidiabetic; immunosuppressive; dermatological;  
KW haemostatic; vulnery; fungicide; antibacterial; virucide; protozoacide;  
KW antiarthritic.  
XX  
XX Homo sapiens.  
XX WO200270539-A2.  
XX  
XX 12-SEP-2002.  
XX  
XX 05-MAR-2002; 2002WO-US05095.  
XX  
XX 05-MAR-2001; 2001US-0799451.  
XX  
XX (HYSE-) HYSEQ INC.  
XX  
XX Tang YT, Zhou P, Goodrich RW, Asundi V, Zhang J, Zhao QA, Ren F;  
PI Xue AJ, Yang Y, Ma Y, Yamazaki V, Chen R, Wang Z, Ghosh M;  
PI Wehrman T, Wang J, Wang D, Drmanac RT;  
XX  
XX WPI; 2002-759812/82.  
DR N-PSDB; ABZ11358.  
XX  
XX New polynucleotides comprising sequences assembled from expressed  
PT sequence tags (ESTs), useful for treating cell-proliferative,  
PT neurodegenerative, autoimmune, genetic, myeloid or lymphoid, or  
PT platelet or coagulation disorders -  
XX

PS Claim 9; SEQ ID NO 1188; 1012pp + Sequence Listing; English.

XX The invention relates to an isolated polynucleotide (I) comprising a  
CC nucleotide sequence selected from any of 948 sequences  
CC (ABZ11119-ABZ12066) or their mature protein coding portion, active domain  
CC coding protein or complementary sequences. The polynucleotides are useful  
CC for identifying expressed genes or for physical mapping of human genome.  
CC The encoded polypeptides (ABP69902-ABP69849) are useful as molecular  
CC weight markers, as a food supplement, for generating antibodies, in  
CC medical imaging, screening and diagnostic assays and for treating  
CC cell-proliferative disorders (cancer), neurodegenerative diseases  
CC (Parkinson's or Alzheimer's disease), autoimmune diseases (multiple  
CC sclerosis, diabetes, lupus) genetic disorders, myeloid or lymphoid  
CC disorders, platelet or coagulation disorders, wound, burns, incision,  
CC ulcers, liver or lung fibrosis, infections (bacterial, viral, fungal,  
CC parasitic), arthritis, etc.

CC Note: The sequence data for this patent did not form part of the printed  
CC specification, but was obtained in electronic format directly from WIPO  
CC at ftp.wipo.int/pub/published\_pct\_sequences.

XX SQ Sequence 97 AA;

Query Match 38.6%; Score 39; DB 23; Length 97;  
Best Local Similarity 53.8%; Pred. No. 1.2e+02;  
Matches 7; Conservative 4; Mismatches 2; Indels 0; Gaps 0;

QY 2 VLMMNIIISKEKK 14  
Db :|||::|||  
14 ILLAMLMVDREKK 26

RESULT 39

AAU78096  
ID AAU78096 standard; Protein; 76 AA.

AC AAU78096;

DT 18-JUN-2002 (first entry)

DE Human DNA binding domain E2F-4.

XX Human; telomerase reverse transcriptase; TERT; Site C; Progeria; burn;  
XX repressor binding site; Hutchinson-Gilford syndrome; AIDS; cancer;  
XX acquired immunodeficiency syndrome; cardiovascular disease; osteoporosis;  
XX skin rejuvenation; immune senescence; bone marrow transplant; skin graft;  
XX neoplastic disease; TERT minimal promoter; DNA binding domain; E2F-4.

OS Homo sapiens.

Key Location/Qualifiers  
FH Domain 7..49  
FT /label= DNA\_binding\_domain

XX WO200216657-A1.

XX 28-FEB-2002.

XX 17-AUG-2001; 2001WO-US25861.

XX 24-AUG-2000; 2000US-227865P.

XX 01-SEP-2000; 2000US-230174P.

XX 05-OCT-2000; 2000US-238345P.

XX (SIER-) SIERRA SCI INC.

XX Andrews WH, Foster CA, Fraser S, Mohammadpour H;

XX WPI; 2002-280952/32.

XX Modulating expression of telomerase reverse transcriptase (TERT) in a  
XX cell, for regulating proliferative capacity of a cell, involves  
XX modulating TERT transcription repression by Site C repressor binding  
XX site -

XX Disclosure; Page 8; 66pp; English.

XX The present invention relates to a new method of modulating expression  
CC of telomerase reverse transcriptase (TERT) from a TERT expression system  
CC that includes a TERT promoter and a Site C repressor binding site. The  
CC method of the invention involves modulating TERT transcription repression  
CC by the Site C repressor binding site. The method of the invention is  
CC useful for modulating expression of TERT for producing a mammalian  
CC antibody. The method is also useful in a variety of different  
CC applications, including immortalisation of cells, production of reagents  
CC for use in life science research, therapeutic applications, and  
CC therapeutic agent screening applications. Increasing TERT expression  
CC delays natural telomeric shortening and/or increases telomeric length and  
CC is useful for treating disease conditions such as Progeria or  
CC Hutchinson-Gilford syndrome, acquired immunodeficiency syndrome (AIDS),  
CC cardiovascular disease, osteoporosis, in skin rejuvenation and to inhibit  
CC immune senescence. The method can be employed to lengthen telomeres of  
CC osteoblast and osteoclast stem cells, encouraging bone replacement and  
CC proper remodeling and reinforcement, and can thus be used in bone marrow  
CC transplants for the treatment of cancer and skin grafts for burn  
CC victims and as such the method improves the survival and effectiveness of  
CC bone marrow and skin cell transplants. Decreasing TERT expression is  
CC useful for treating cellular proliferative disease conditions, including  
CC neoplastic disease conditions e.g. cancer. The present amino acid  
CC sequence represents the human DNA binding domain E2F-4.

XX SQ Sequence 76 AA;

Query Match 38.1%; Score 38.5; DB 23; Length 76;  
Best Local Similarity 42.9%; Pred. No. 1.1e+02;  
Matches 9; Conservative 4; Mismatches 7; Indels 1; Gaps 1;

QY 1 NVLMAMNIIISKE-KKEIKWIG 20  
Db :|||::|||  
53 NVLEGIGLIEKKSNSIQWKG 73

RESULT 40

ABB82987

ID ABB82987 standard; Protein; 76 AA.

AC ABB82987;

XX 14-APR-2003 (first entry)

DE Repressor protein E2F-4 DNA binding domain.

XX Telomerase reverse transcriptase; TERT; Site C repressor; transcription;  
XX cytosolic; immunostimulant; anti-HIV; vulnery; telomerase; human;  
XX repressor protein; E2F-4.

OS Homo sapiens.

XX WO2002101010-A2.

XX 19-DEC-2002.

XX 06-JUN-2002; 2002WO-US17959.

XX 07-JUN-2001; 2001US-296992P.

XX (SIER-) SIERRA SCI INC.

XX Foster CA, Fraser S, Mohammadpour H, Andrews WH;

XX WPI; 2003-167401/16.

XX Modulating expression of telomerase reverse transcriptase TERT by  
XX blocking repression of TERT transcription, useful for the diagnosis and  
XX treatment of disorders associated with aberrant telomerase activity  
XX such as cancer and HIV -

PS Disclosure; Page 7; 47pp; English.

XX The invention relates to modulating expression of telomerase reverse transcriptase (TERT) expression system that includes a TERT promoter and a Site C repressor binding site. The method involves modulating TERT transcription repression by the Site C repressor binding site. The methods and compositions of the present invention are useful for the immortalization of cells, production of reagents in life science research, therapeutic agent screening applications, diagnosis and treatment of disorders associated with aberrant telomerase activity such as cancer, progeria, immune senescence, HIV, and in skin rejuvenation. The present sequence represents the DNA binding domain of a repressor protein E2F-4.

XX SQ Sequence 76 AA;

Query Match 38.1%; Score 38.5; DB 24; Length 76;  
Best Local Similarity 42.3%; Pred. No. 1.1e+02;  
Matches 9; Conservative 4; Mismatches 7; Indels 1; Gaps 1;

QY 1 NVLMANNIISKE-KKEIKWIG 20  
Db 53 NVLEGIGLIEKSKNSIQWKG 73

RESULT 41  
ABG54577  
ID ABG54577 standard; Peptide; 31 AA.  
AC ABG54577;  
XX  
DT 25-FEB-2003 (first entry)  
XX  
DE Human liver peptide, SEQ ID No 33225.  
XX  
XX Human; liver; cirrhosis; hyperlipoproteinaemia; hyperlipidaemia;  
KW hypercholesterolaemia; coronary heart disease.  
XX  
OS Homo sapiens.  
XX  
PN WO200157273-A2.  
XX  
PD 09-AUG-2001.  
XX  
PF 30-JAN-2001; 2001WO-US00664.  
XX  
PR 04-FEB-2000; 2000US-0180312.  
PR 26-MAY-2000; 2000US-0207456.  
PR 30-JUN-2000; 2000US-0608408.  
PR 03-AUG-2000; 2000US-0632366.  
PR 21-SEP-2000; 2000US-0234687.  
PR 27-SEP-2000; 2000US-0236359.  
PR 04-OCT-2000; 2000GB-0024263.  
XX  
PA (MOLE-) MOLECULAR DYNAMICS INC.  
XX  
PI Penn SG, Hanzel DK, Chen W, Rank DR;  
XX  
DR WPI; 2001-48898/53.  
XX  
PT Human genome-derived single exon nucleic acid probes useful for  
PT analysing gene expression in human adult liver -  
XX  
PS Claim 27; SEQ ID No 33225; 658pp; English.  
XX  
CC The invention relates to a single exon nucleic acid probe (SENP) (I) for  
CC measuring human gene expression in a sample derived from human adult  
CC liver, comprising one of 13109 defined nucleotide sequences given in the  
CC specification (or complements/ fragments). The probe hybridises at high  
CC stringency to a nucleic acid molecule expressed in the human adult  
CC liver. (I) may be used for predicting, measuring and displaying gene  
CC expression in samples derived from human adult liver. The genes  
CC identified may be involved in genetic liver diseases such as cirrhosis,

CC hyperlipoproteinaemia, hyperlipidaemia and hypercholesterolaemia which  
CC is associated with coronary heart disease. ABG47348-ABG59930 represent  
CC human liver single exon encoded peptides of the invention.  
CC Note: The sequence information for this patent does not appear in the  
CC printed specification but was obtained in electronic format directly  
CC from WIPO at fcp.wipo.int/pub/published\_pct\_sequences.

XX SQ Sequence 31 AA;

Query Match 37.6%; Score 38; DB 22; Length 31;  
Best Local Similarity 50.0%; Pred. No. 47;  
Matches 6; Conservative 4; Mismatches 2; Indels 0; Gaps 0;

QY 7 NIISKEKEIKW 18  
Db 7 NMISKQRKKW 18

RESULT 42  
ABB39547  
ID ABB39547 standard; Peptide; 31 AA.  
XX  
AC ABB39547;  
XX  
DT 04-FEB-2002 (first entry)  
XX  
DE Peptide #7053 encoded by human foetal liver single exon probe.  
XX  
KW Human; foetal liver; gene expression; single exon nucleic acid probe.  
XX  
OS Homo sapiens.  
XX  
PN WO200157277-A2.  
XX  
PD 09-AUG-2001.  
XX  
PF 30-JAN-2001; 2001WO-US00669.  
XX  
PR 04-FEB-2000; 2000US-0180312.  
PR 26-MAY-2000; 2000US-0207456.  
PR 30-JUN-2000; 2000US-0608408.  
PR 03-AUG-2000; 2000US-0632366.  
PR 21-SEP-2000; 2000US-0234687.  
PR 27-SEP-2000; 2000US-0236359.  
PR 04-OCT-2000; 2000GB-0024263.  
XX  
PA (MOLE-) MOLECULAR DYNAMICS INC.  
XX  
PI Penn SG, Hanzel DK, Chen W, Rank DR;  
XX  
DR WPI; 2001-483447/52.  
XX  
PT Human genome-derived single exon nucleic acid probes useful for  
PT analysing gene expression in human fetal liver -  
XX  
PS Claim 27; SEQ ID No 32182; 639pp + sequence listing; English.  
XX  
CC The invention relates to a single exon nucleic acid probe for  
CC measuring human gene expression in a sample derived from human foetal  
CC liver. The single exon nucleic acid probes may be used for predicting,  
CC measuring and displaying gene expression in samples derived from human  
CC fetal liver. The present sequence is a peptide encoded by a single exon  
CC nucleic acid probe of the invention.  
CC Note: The sequence data for this patent did not form part of the  
CC printed specification, but was obtained in electronic format directly  
CC from WIPO at fcp.wipo.int/pub/published\_pct\_sequences.

XX SQ Sequence 31 AA;

Query Match 37.6%; Score 38; DB 22; Length 31;  
Best Local Similarity 50.0%; Pred. No. 47;  
Matches 6; Conservative 4; Mismatches 2; Indels 0; Gaps 0;





```

PR 30-JUN-2000; 2000US-0608408.
PR 03-AUG-2000; 2000US-0632366.
PR 21-SEP-2000; 2000US-0234687.
PR 27-SEP-2000; 2000US-0236359.
PR 04-OCT-2000; 2000GB-0024263.
XX
XX (MOLE-) MOLECULAR DYNAMICS INC.
XX
XX Penn SG, Hanzel DK, Chen W, Rank DR;
XX WPI; 2001-488897/53.
XX
XX Human genome-derived single exon nucleic acid probes useful for
XX analyzing gene expression in human placenta -
XX
XX Claim 27; SEQ ID No 33370; 654pp; English.
XX
XX The present invention relates to single exon nucleic acid probes (SENP:
XX see AA13135-AA15746). The present sequence is a peptide encoded by one
XX such probe. The probes are useful for producing a microarray for
XX predicting, measuring and displaying gene expression in samples derived
XX from human placenta. The probes are useful for antenatal diagnosis of
XX human genetic disorders.
XX
XX Sequence 31 AA;
SQ
Query Match 37.6%; Score 38; DB 22; Length 31;
Best Local Similarity 50.0%; Pred. No. 47;
Matches 6; Conservative 4; Mismatches 2; Indels 0; Gaps 0;
QY 7 NIISKEKEIKW 18
|:|:|:|:|:|
DB 7 NMISQSRKKW 18

```

Search completed: February 17, 2004, 10:53:42  
Job time : 39.0198 secs

**THIS PAGE BLANK (USPTO)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: February 17, 2004, 10:53:48 ; Search time 29.3069 Seconds  
(without alignments)  
142.889 Million cell updates/sec

Title: US-09-900-147-4

Perfect score: 101

Sequence: 1 NVLMANNIISKEKEIKWIG 20

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 801455 seqs, 209382283 residues

Total number of hits satisfying chosen parameters: 318354

Minimum DB seq length: 0

Maximum DB seq length: 100

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications\_AA:

1: /cgn2\_6/ptodata/1/pubpaa/US07\_PUBCOMB.pep.\*  
2: /cgn2\_6/ptodata/1/pubpaa/PCT\_NEW\_PUB.pep.\*  
3: /cgn2\_6/ptodata/1/pubpaa/US06\_NEW\_PUB.pep.\*  
4: /cgn2\_6/ptodata/1/pubpaa/US06\_PUBCOMB.pep.\*  
5: /cgn2\_6/ptodata/1/pubpaa/US07\_NEW\_PUB.pep.\*  
6: /cgn2\_6/ptodata/1/pubpaa/PCTUS\_PUBCOMB.pep.\*  
7: /cgn2\_6/ptodata/1/pubpaa/US08\_NEW\_PUB.pep.\*  
8: /cgn2\_6/ptodata/1/pubpaa/US08\_PUBCOMB.pep.\*  
9: /cgn2\_6/ptodata/1/pubpaa/US09A\_PUBCOMB.pep.\*  
10: /cgn2\_6/ptodata/1/pubpaa/US09B\_PUBCOMB.pep.\*  
11: /cgn2\_6/ptodata/1/pubpaa/US09C\_PUBCOMB.pep.\*  
12: /cgn2\_6/ptodata/1/pubpaa/US09\_NEW\_PUB.pep.\*  
13: /cgn2\_6/ptodata/1/pubpaa/US10A\_PUBCOMB.pep.\*  
14: /cgn2\_6/ptodata/1/pubpaa/US10B\_PUBCOMB.pep.\*  
15: /cgn2\_6/ptodata/1/pubpaa/US10C\_PUBCOMB.pep.\*  
16: /cgn2\_6/ptodata/1/pubpaa/US10\_NEW\_PUB.pep.\*  
17: /cgn2\_6/ptodata/1/pubpaa/US60\_NEW\_PUB.pep.\*  
18: /cgn2\_6/ptodata/1/pubpaa/US60\_PUBCOMB.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	101	100.0	20	10	US-09-900-147-4
2	101	100.0	30	10	US-09-900-147-6
3	101	100.0	37	10	US-09-900-147-1
4	101	100.0	74	15	US-10-214-188-10
5	51	50.5	19	10	US-09-900-147-3
6	51	50.5	19	10	US-09-900-147-16
7	50	49.5	15	10	US-09-900-147-10
8	50	49.5	29	10	US-09-764-877-1380
9	50	49.5	29	12	US-10-242-515-1380
10	47	46.5	19	10	US-09-900-147-15
11	46	45.5	16	10	US-09-900-147-5
12	45.5	45.0	74	15	US-10-214-188-6
13	45	44.6	19	10	US-09-900-147-17
14	44.5	44.1	74	15	US-10-214-188-5
15	44.5	44.1	85	11	US-09-932-581-5

16	44.5	44.1	85	12	US-10-338-294-5	Sequence 5, Appli
17	44.5	44.1	85	14	US-10-165-614-2	Sequence 2, Appli
18	42.5	42.1	74	15	US-10-214-188-7	Sequence 7, Appli
19	42	41.6	9	10	US-09-900-147-2	Sequence 2, Appli
20	40	39.6	57	15	US-10-097-065-532	Sequence 532, App
21	40	39.6	63	12	US-09-864-408A-3370	Sequence 3370, Ap
22	40	39.6	75	9	US-09-864-761-35555	Sequence 35555, A
23	39.5	39.1	75	15	US-10-214-188-9	Sequence 9, Appli
24	39	38.6	84	12	US-10-407-920-19	Sequence 19, Appli
25	38.5	38.1	76	11	US-09-932-581-6	Sequence 6, Appli
26	38.5	38.1	76	12	US-10-338-294-6	Sequence 6, Appli
27	38.5	38.1	76	12	US-10-029-386-29071	Sequence 29071, A
28	38.5	38.1	76	14	US-10-165-614-3	Sequence 3, Appli
29	38	37.6	31	9	US-09-864-761-43984	Sequence 43984, A
30	38	37.6	69	15	US-10-083-357-694	Sequence 694, App
31	38	37.6	81	12	US-10-029-386-30735	Sequence 30735, A
32	38	37.6	85	9	US-09-916-790-25	Sequence 25, Appl
33	38	37.6	100	9	US-09-922-138-20	Sequence 20, Appl
34	35.5	35.1	77	9	US-09-867-550-1920	Sequence 1920, Ap
35	35	34.7	58	9	US-09-864-761-33458	Sequence 33458, A
36	35	34.7	98	12	US-09-864-408A-4500	Sequence 4500, Ap
37	34	33.7	14	10	US-09-900-147-11	Sequence 11, Appl
38	34	33.7	18	15	US-10-157-031-192	Sequence 192, App
39	34	33.7	32	9	US-09-864-761-48415	Sequence 48415, A
40	34	33.7	35	9	US-09-764-869-723	Sequence 723, App
41	34	33.7	35	12	US-10-227-577-723	Sequence 723, App
42	34	33.7	35	15	US-10-091-504-723	Sequence 723, App
43	34	33.7	49	9	US-09-864-761-47714	Sequence 47714, A
44	34	33.7	56	12	US-10-029-386-33245	Sequence 33245, A
45	34	33.7	59	9	US-09-764-869-845	Sequence 845, App

#### ALIGNMENTS

##### RESULT 1

US-09-900-147-4  
; Sequence 4, Application US/09900147  
; Patent No. US20020103121A1  
; GENERAL INFORMATION:  
; APPLICANT: La Thangue, Nicholas B  
; APPLICANT: Bandara, Lasantha R  
; TITLE OF INVENTION: Peptide antagonists of DP transcription factors  
; FILE REFERENCE: 620-67  
; CURRENT APPLICATION NUMBER: US/09/900,147  
; CURRENT FILING DATE: 2001-07-09  
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 09/308,935  
; PRIOR FILING DATE: EARLIER FILING DATE: 1993-05-27  
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: GB 962589.7  
; PRIOR FILING DATE: EARLIER FILING DATE: 1996-12-20  
; NUMBER OF SEQ ID NOS: 18  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 4  
; LENGTH: 20  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Description of Artificial Sequence: Synthetic  
US-09-900-147-4

Query Match 100.0%; Score 101; DB 10; Length 20;  
Best Local Similarity 100.0%; Pred. No. 1.3e-08;  
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 NVLMANNIISKEKEIKWIG 20

Db 1 NVLMANNIISKEKEIKWIG 20

##### RESULT 2

US-09-900-147-6  
; Sequence 6, Application US/09900147  
; Patent No. US20020103121A1

```

; GENERAL INFORMATION:
; APPLICANT: La Thangue, Nicholas B
; APPLICANT: Bandara, Lasantha R
; TITLE OF INVENTION: Peptide antagonists of DP transcription factors
; FILE REFERENCE: 620-67
; CURRENT APPLICATION NUMBER: US/09/900,147
; CURRENT FILING DATE: 2001-07-09
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 09/308,935
; PRIOR FILING DATE: EARLIER FILING DATE: 1999-05-27
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: GB 9626589.7
; PRIOR FILING DATE: EARLIER FILING DATE: 1996-12-20
; NUMBER OF SEQ ID NOS: 18
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 6
; LENGTH: 30
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Synthetic
; US-09-900-147-6

Query Match 100.0%; Score 101; DB 10; Length 30;
Best Local Similarity 100.0%; Pred. No. 2e-08;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 NVLMAMNIISKEKEIKWIG 20
Db 5 NVLMAMNIISKEKEIKWIG 24

RESULT 3
US-09-900-147-1
; Sequence 1, Application US/09900147
; Patent No. US20020103121A1
; GENERAL INFORMATION:
; APPLICANT: La Thangue, Nicholas B
; APPLICANT: Bandara, Lasantha R
; TITLE OF INVENTION: Peptide antagonists of DP transcription factors
; FILE REFERENCE: 620-67
; CURRENT APPLICATION NUMBER: US/09/900,147
; CURRENT FILING DATE: 2001-07-09
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 09/308,935
; PRIOR FILING DATE: EARLIER FILING DATE: 1999-05-27
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: GB 9626589.7
; PRIOR FILING DATE: EARLIER FILING DATE: 1996-12-20
; NUMBER OF SEQ ID NOS: 18
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1
; LENGTH: 37
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Synthetic
; US-09-900-147-1

Query Match 100.0%; Score 101; DB 10; Length 37;
Best Local Similarity 100.0%; Pred. No. 2.5e-08;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 NVLMAMNIISKEKEIKWIG 20
Db 12 NVLMAMNIISKEKEIKWIG 31

RESULT 4
US-10-214-188-10
; Sequence 10, Application US/10214188
; Publication No. US2003002260A1
; GENERAL INFORMATION:
; APPLICANT: LA THANGUE, NICHOLAS B.
; APPLICANT: BERNARDS, RENE
; APPLICANT: HJLMANS, ELEANORE M.
; TITLE OF INVENTION: TRANSCRIPTION FACTOR E2F-5

; NUMBER OF SEQUENCES: 25
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: NIXON & VANDERHYE P.C.
; STREET: 1100 NORTH GLEBE ROAD
; CITY: ARLINGTON
; STATE: VIRGINIA
; COUNTRY: U.S.A.
; ZIP: 22201-4714
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/10/214,188
; FILING DATE: 08-Aug-2002
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/894,139
; FILING DATE: 13-AUG-1997
; ATTORNEY/AGENT INFORMATION:
; NAME: WILSON, MARY J.
; REGISTRATION NUMBER: 32,955
; REFERENCE/DOCKET NUMBER: 620-22
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703) 816-4000
; TELEFAX: (703) 816-4100
; INFORMATION FOR SEQ ID NO: 10:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 74 amino acids
; TYPE: amino acid
; STRANDEDNESS: <Unknown>
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; SEQUENCE DESCRIPTION: SEQ ID NO: 10:
; US-10-214-188-10

Query Match 100.0%; Score 101; DB 15; Length 74;
Best Local Similarity 100.0%; Pred. No. 5.1e-08;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 NVLMAMNIISKEKEIKWIG 20
Db 54 NVLMAMNIISKEKEIKWIG 73

RESULT 5
US-09-900-147-3
; Sequence 3, Application US/09900147
; Patent No. US20020103121A1
; GENERAL INFORMATION:
; APPLICANT: La Thangue, Nicholas B
; APPLICANT: Bandara, Lasantha R
; TITLE OF INVENTION: Peptide antagonists of DP transcription factors
; FILE REFERENCE: 620-67
; CURRENT APPLICATION NUMBER: US/09/900,147
; CURRENT FILING DATE: 2001-07-09
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 09/308,935
; PRIOR FILING DATE: EARLIER FILING DATE: 1999-05-27
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: GB 9626589.7
; PRIOR FILING DATE: EARLIER FILING DATE: 1996-12-20
; NUMBER OF SEQ ID NOS: 19
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 3
; LENGTH: 19
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Synthetic
; US-09-900-147-3

Query Match 50.5%; Score 51; DB 10; Length 19;
Best Local Similarity 100.0%; Pred. No. 0.49;
```

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: February 17, 2004, 10:50:13 ; Search time 14.4554 Seconds  
(without alignments)  
58.540 Million cell updates/sec

Title: US-09-900-147-4

Perfect score: 101

Sequence: 1 NVLMANNIISKEKEIKWIG 20

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 328717 seqs, 42310858 residues

Total number of hits satisfying chosen parameters: 228043

Minimum DB seq length: 0

Maximum DB seq length: 100

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Issued Patents AA.\*

1: /cgn2\_6/ptodata/1/iaa/5A COMB.pcp.\*

2: /cgn2\_6/ptodata/1/iaa/5B COMB.pcp.\*

3: /cgn2\_6/ptodata/1/iaa/6A COMB.pcp.\*

4: /cgn2\_6/ptodata/1/iaa/6B COMB.pcp.\*

5: /cgn2\_6/ptodata/1/iaa/PCTUS COMB.pcp.\*

6: /cgn2\_6/ptodata/1/iaa/backfiles1.pcp.\*

pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	101	100.0	20	3	US-09-308-935-4
2	101	100.0	30	3	US-09-308-935-6
3	101	100.0	37	3	US-09-308-935-1
4	101	100.0	72	2	US-08-428-131-11
5	101	100.0	72	3	US-09-078-596-11
6	101	100.0	74	4	US-08-894-139-10
7	51	50.5	19	3	US-09-308-935-3
8	51	50.5	19	3	US-09-308-935-16
9	50	49.5	15	3	US-09-308-935-10
10	47	46.5	15	3	US-09-308-935-15
11	46	45.5	16	3	US-09-308-935-5
12	46	45.5	17	2	US-08-428-131-13
13	46	45.5	17	3	US-09-078-596-13
14	45.5	45.0	74	4	US-08-894-139-6
15	45	44.6	19	3	US-09-308-935-17
16	44.5	44.1	73	2	US-08-428-131-12
17	44.5	44.1	73	3	US-09-078-596-12
18	44.5	44.1	74	4	US-08-894-139-5
19	42.5	42.1	74	4	US-08-894-139-7
20	42	41.6	9	3	US-09-308-935-2
21	39.5	39.1	75	4	US-08-894-139-9
22	39	38.6	78	4	US-09-328-352-7148
23	37	36.6	87	4	US-09-107-532A-5431
24	35	34.7	26	1	US-07-948-357-3
25	35	34.7	26	3	US-08-450-417-3
26	35	34.7	26	3	US-08-449-741-3
27	35	34.7	26	3	US-08-782-997A-5

28 35 34.7 26 3 US-08-449-754-3 Sequence 3, Appli  
29 35 34.7 26 3 US-09-465-646-3 Sequence 3, Appli  
30 35 34.7 36 1 US-07-948-357-4 Sequence 4, Appli  
31 35 34.7 36 2 US-08-450-417-4 Sequence 4, Appli  
32 35 34.7 36 3 US-08-449-741-4 Sequence 4, Appli  
33 35 34.7 36 3 US-08-449-754-4 Sequence 4, Appli  
34 35 34.7 36 3 US-09-465-646-4 Sequence 14, Appli  
35 35 34.7 63 1 US-08-194-338-14 Sequence 14, Appli  
36 34 33.7 21 4 US-09-308-935-11 Sequence 15, Appli  
37 34 33.7 21 4 US-09-230-548-15 Sequence 15, Appli  
38 34 33.7 26 4 US-09-230-548-26 Sequence 26, Appli  
39 34 33.7 37 4 US-09-230-548-31 Sequence 31, Appli  
40 34 33.7 46 4 US-09-230-548-25 Sequence 25, Appli  
41 33.5 33.2 69 4 US-08-894-139-8 Sequence 8, Appli  
42 33.5 33.2 86 2 US-08-343-443B-8 Sequence 8, Appli  
43 33 32.7 61 3 US-08-743-975-9 Sequence 9, Appli  
44 33 32.7 61 4 US-09-263-811-9 Sequence 9, Appli  
45 33 32.7 66 4 US-09-205-258-381 Sequence 381, App

#### ALIGNMENTS

RESULT 1  
US-09-308-935-4  
; Sequence 4, Application US/09308935  
; Patent No. 6268334  
; GENERAL INFORMATION:  
; APPLICANT: La Thangue, Nicholas B  
; APPLICANT: Bandara, Lasantha R  
; TITLE OF INVENTION: Peptide antagonists of DP transcription factors  
; FILE REFERENCE: 620-67  
; CURRENT APPLICATION NUMBER: US/09/308,935  
; CURRENT FILING DATE: 1999-05-27  
; EARLIER APPLICATION NUMBER: PCT/GB97/03506  
; EARLIER FILING DATE: 1997-12-22  
; EARLIER APPLICATION NUMBER: GB 9626589.7  
; EARLIER FILING DATE: 1996-12-20  
; NUMBER OF SEQ ID NOS: 18  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 4  
; LENGTH: 20  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Description of Artificial Sequence: Synthetic  
US-09-308-935-4

Query Match 100.0%; Score 101; DB 3; Length 20;  
Best Local Similarity 100.0%; Pred. No. 2.3e-09;  
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 NVLMANNIISKEKEIKWIG 20  
Db 1 NVLMANNIISKEKEIKWIG 20

RESULT 2  
US-09-308-935-6  
; Sequence 6, Application US/09308935  
; Patent No. 6268334  
; GENERAL INFORMATION:  
; APPLICANT: La Thangue, Nicholas B  
; APPLICANT: Bandara, Lasantha R  
; TITLE OF INVENTION: Peptide antagonists of DP transcription factors  
; FILE REFERENCE: 620-67  
; CURRENT APPLICATION NUMBER: US/09/308,935  
; CURRENT FILING DATE: 1999-05-27  
; EARLIER APPLICATION NUMBER: PCT/GB97/03506  
; EARLIER FILING DATE: 1997-12-22  
; EARLIER APPLICATION NUMBER: GB 9626589.7  
; EARLIER FILING DATE: 1996-12-20  
; NUMBER OF SEQ ID NOS: 18

2

Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 NVLMAMNIISKEKEIKWIG 20  
| | | | | | | | | | | | | | | | | |  
Db 15 NVLMAMNIISKEKEIKWIG 34

## RESULT 6

US-08-894-139-10  
; Sequence 10, Application US/08894139  
; Patent No. 6448376

; GENERAL INFORMATION:

; APPLICANT: LA THANGUE, NICHOLAS B.

; APPLICANT: BERNARDS, RENE

; APPLICANT: HIJMAN, ELEANORE M.

; TITLE OF INVENTION: TRANSCRIPTION FACTOR E2F-5

; NUMBER OF SEQUENCES: 25

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: NIXON & VANDERHYE P.C.

; STREET: 1100 NORTH GLEBE ROAD

; CITY: ARLINGTON

; STATE: VIRGINIA

; COUNTRY: U.S.A.

; ZIP: 22201-4714

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: Patent In Release #1.0, Version #1.30

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/894,139

; FILING DATE: 13-AUG-1997

; CLASSIFICATION: 536

; ATTORNEY/AGENT INFORMATION:

; NAME: WILSON, MARY J.

; REGISTRATION NUMBER: 32,955

; REFERENCE/DOCKET NUMBER: 620-22

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: (703) 816-4000

; TELEFAX: (703) 816-4100

; INFORMATION FOR SEQ ID NO: 10:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 74 amino acids

; TYPE: amino acid

; STRANDEDNESS:

; TOPOLOGY: linear

; MOLECULE TYPE: peptide

US-08-894-139-10

Query Match 100.0%; Score 101; DB 4; Length 74;  
Best Local Similarity 100.0%; Pred. No. 9.4e-09;  
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 NVLMAMNIISKEKEIKWIG 20  
| | | | | | | | | | | | | | | | | |  
Db 54 NVLMAMNIISKEKEIKWIG 73

## RESULT 7

US-09-308-935-3

; Sequence 3, Application US/09308935

; Patent No. 6268334

; GENERAL INFORMATION:

; APPLICANT: La Thangue, Nicholas B

; APPLICANT: Bandara, Lasantha R

; TITLE OF INVENTION: Peptide antagonists of DP transcription factors

; FILE REFERENCE: 620-67

; CURRENT APPLICATION NUMBER: US/09/308,935

; EARLIER FILING DATE: 1999-05-27

; EARLIER APPLICATION NUMBER: PCT/GB97/03506

; EARLIER FILING DATE: 1997-12-22

; EARLIER APPLICATION NUMBER: GB 9626589.7

; EARLIER FILING DATE: 1996-12-20

; TYPE: PRT

US-09-308-935-3

Query Match 100.0%; Score 101; DB 4; Length 74;  
Best Local Similarity 100.0%; Pred. No. 9.4e-09;  
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 NVLMAMNIISKEKEIKWIG 20  
| | | | | | | | | | | | | | | | | |  
Db 54 NVLMAMNIISKEKEIKWIG 73

; NUMBER OF SEQ ID NOS: 18  
; SOFTWARE: Patent in Ver. 2.1  
; SEQ ID NO 3  
; LENGTH: 19  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Description of Artificial Sequence: Synthetic  
US-09-308-935-3

Query Match 50.5%; Score 51; DB 3; Length 19;  
Best Local Similarity 100.0%; Pred. No. 0.12;  
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 NVLMAMNIISK 11  
| | | | | | | | | | | | | | | | | |  
Db 9 NVLMAMNIISK 19

## RESULT 8

US-09-308-935-16

; Sequence 16, Application US/09308935

; Patent No. 6268334

; GENERAL INFORMATION:

; APPLICANT: La Thangue, Nicholas B

; APPLICANT: Bandara, Lasantha R

; TITLE OF INVENTION: Peptide antagonists of DP transcription factors

; FILE REFERENCE: 620-67

; CURRENT APPLICATION NUMBER: US/09/308,935

; CURRENT FILING DATE: 1999-05-27

; EARLIER APPLICATION NUMBER: PCT/GB97/03506

; EARLIER FILING DATE: 1997-12-22

; EARLIER APPLICATION NUMBER: GB 9626589.7

; EARLIER FILING DATE: 1996-12-20

; NUMBER OF SEQ ID NOS: 18

; SOFTWARE: Patent in Ver. 2.1

; SEQ ID NO 16

; LENGTH: 19

; TYPE: PRT

; ORGANISM: Artificial Sequence

; FEATURE:

; OTHER INFORMATION: Description of Artificial Sequence: Mutant peptide

US-09-308-935-16

Query Match 50.5%; Score 51; DB 3; Length 19;  
Best Local Similarity 100.0%; Pred. No. 0.12;  
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 NVLMAMNIISK 11  
| | | | | | | | | | | | | | | | | |  
Db 9 NVLMAMNIISK 19

## RESULT 9

US-09-308-935-10

; Sequence 10, Application US/09308935

; Patent No. 6268334

; GENERAL INFORMATION:

; APPLICANT: La Thangue, Nicholas B

; APPLICANT: Bandara, Lasantha R

; TITLE OF INVENTION: Peptide antagonists of DP transcription factors

; FILE REFERENCE: 620-67

; CURRENT APPLICATION NUMBER: US/09/308,935

; CURRENT FILING DATE: 1999-05-27

; EARLIER APPLICATION NUMBER: PCT/GB97/03506

; EARLIER FILING DATE: 1997-12-22

; EARLIER APPLICATION NUMBER: GB 9626589.7

; EARLIER FILING DATE: 1996-12-20

; NUMBER OF SEQ ID NOS: 18

; SOFTWARE: Patent in Ver. 2.1

; SEQ ID NO 10

; LENGTH: 15

; TYPE: PRT

US-09-308-935-10

Query Match 50.5%; Score 51; DB 3; Length 19;  
Best Local Similarity 100.0%; Pred. No. 0.12;  
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 NVLMAMNIISK 11  
| | | | | | | | | | | | | | | | | |  
Db 9 NVLMAMNIISK 19

Query Match	45.5%; Score 46; DB 3; Length 16;
Best Local Similarity	100.0%; Pred. No. 0.59;
Matches	10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy	1 NVLMAMNIIS 10
Db	7 NVLMAMNIIS 16
RESULT 12	
US-08-428-131-13	
Sequence 13, Application US/08428131	
Patent No. 5863757	
GENERAL INFORMATION:	
APPLICANT: La Thangue, Nicholas Barrie	
TITLE OF INVENTION: Transcription Factor DP-1	
NUMBER OF SEQUENCES: 14	
CORRESPONDENCE ADDRESS:	
ADDRESSEE: Nixon & Vanderhye	
STREET: 1100 No. 5863757th Glebe Road, 8th Floor	
CITY: Arlington	
STATE: Virginia	
COUNTRY: U.S.A.	
ZIP: 22201-4714	
COMPUTER READABLE FORM:	
MEDIUM TYPE: Floppy disk	
COMPUTER: IBM PC compatible	
OPERATING SYSTEM: PC-DOS/MS-DOS	
SOFTWARE: Patentin Release #1.0, Version #1.25 (EPO)	
CURRENT APPLICATION NUMBER: US/08/428,131	
FILING DATE: 23-JUN-1995	
CLASSIFICATION: 514	
ATTORNEY/AGENT INFORMATION:	
NAME: Arthur R. Crawford	
REGISTRATION NUMBER: 25,327	
REFERENCE/DOCKET NUMBER: 117-181	
TELECOMMUNICATION INFORMATION:	
TELEPHONE: (703) 816-4000	
TELEFAX: (703) 816-4100	
INFORMATION FOR SEQ ID NO: 13:	
SEQUENCE CHARACTERISTICS:	
LENGTH: 17 amino acids	
TYPE: amino acid	
STRANDEDNESS: single	
TOPOLOGY: linear	
MOLECULE TYPE: protein	
US-08-428-131-13	
Query Match	45.5%; Score 46; DB 2; Length 17;
Best Local Similarity	100.0%; Pred. No. 0.63;
Matches	10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy	1 NVLMAMNIIS 10
Db	8 NVLMAMNIIS 17
RESULT 13	
US-09-078-596-13	
Sequence 13, Application US/09078596	
Patent No. 6150116	
GENERAL INFORMATION:	
APPLICANT: La Thangue, Nicholas Barrie	
TITLE OF INVENTION: Transcription Factor DP-1	
NUMBER OF SEQUENCES: 14	
CORRESPONDENCE ADDRESS:	
ADDRESSEE: Nixon & Vanderhye	
STREET: 1100 No. 6150116th Glebe Road, 8th Floor	
CITY: Arlington	
STATE: Virginia	
COUNTRY: U.S.A.	
ZIP: 22201-4714	



```
/
/ COMPUTER READABLE FORM:
/ MEDIUM TYPE: Floppy disk
/ COMPUTER: IBM PC compatible
/ OPERATING SYSTEM: PC-DOS/MS-DOS
/ SOFTWARE: PatentIn Release #1.0, Version #1.25 (BPO)
/ CURRENT APPLICATION DATA:
/ APPLICATION NUMBER: US/09/078,596
/ FILING DATE:
/ CLASSIFICATION:
/ PRIOR APPLICATION DATA:
/ APPLICATION NUMBER: US/08/428,131
/ FILING DATE: 23-JUN-1995
/ ATTORNEY/AGENT INFORMATION:
/ NAME: Arthur R. Crawford
/ REGISTRATION NUMBER: 25,327
/ REFERENCE/DOCKET NUMBER: 117-181
/ TELECOMMUNICATION INFORMATION:
/ TELEPHONE: (703) 816-4000
/ TELEFAX: (703) 816-4100
/ INFORMATION FOR SEQ ID NO: 13:
/ SEQUENCE CHARACTERISTICS:
/ LENGTH: 17 amino acids
/ TYPE: amino acid
/ STRANDEDNESS: single
/ TOPOLOGY: linear
/ MOLECULE TYPE: protein
/ US-09-078-596-13

Query Match 45.5%; Score 46; DB 3; Length 17;
Best Local Similarity 100.0%; Pred. No. 0.63;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 NVLMAMNIIS 10
Db 8 NVLMAMNIIS 17

RESULT 14
US-08-894-139-6
/ Sequence 6, Application US/08894139
/ Patent No. 6448376
/ GENERAL INFORMATION:
/ APPLICANT: LA THANGUE, NICHOLAS B.
/ APPLICANT: BERNARDS, RENE
/ APPLICANT: HIJWANS, ELEANORE M.
/ TITLE OF INVENTION: TRANSCRIPTION FACTOR E2F-5
/ NUMBER OF SEQUENCES: 25
/ CORRESPONDENCE ADDRESS:
/ ADDRESSEE: NIXON & VANDERHYE P.C.
/ STREET: 1100 NORTH GLEBE ROAD
/ CITY: ARLINGTON
/ STATE: VIRGINIA
/ COUNTRY: U.S.A.
/ ZIP: 22201-4714
/ COMPUTER READABLE FORM:
/ MEDIUM TYPE: Floppy disk
/ COMPUTER: IBM PC compatible
/ OPERATING SYSTEM: PC-DOS/MS-DOS
/ SOFTWARE: PatentIn Release #1.0, Version #1.30
/ CURRENT APPLICATION DATA:
/ APPLICATION NUMBER: US/08/894,139
/ FILING DATE: 13-AUG-1997
/ CLASSIFICATION: 536
/ ATTORNEY/AGENT INFORMATION:
/ NAME: WILSON, MARY J.
/ REGISTRATION NUMBER: 32,955
/ REFERENCE/DOCKET NUMBER: 620-22
/ TELECOMMUNICATION INFORMATION:
/ TELEPHONE: (703) 816-4000
/ TELEFAX: (703) 816-4100
/ INFORMATION FOR SEQ ID NO: 6:
/ SEQUENCE CHARACTERISTICS:
/ LENGTH: 74 amino acids

/
/ TYPE: amino acid
/ STRANDEDNESS:
/ TOPOLOGY: linear
/ MOLECULE TYPE: peptide
/ US-08-894-139-6

Query Match 45.0%; Score 45.5; DB 4; Length 74;
Best Local Similarity 42.9%; Pred. No. 3.7;
Matches 9; Conservative 6; Mismatches 5; Indels 1; Gaps 1;

QY 1 NVLMAMNIISKKK-EIKWIG 20
Db 53 NVLEGIQLIRKKRKNHIQWVG 73

RESULT 15
US-09-308-935-17
/ Sequence 17, Application US/09308935
/ Patent No. 6268334
/ GENERAL INFORMATION:
/ APPLICANT: La Thangue, Nicholas B
/ APPLICANT: Bandara, Lasantha R
/ TITLE OF INVENTION: Peptide antagonists of DP transcription factors
/ FILE REFERENCE: 620-67
/ CURRENT APPLICATION NUMBER: US/09/308,935
/ CURRENT FILING DATE: 1999-05-27
/ EARLIER APPLICATION NUMBER: PCT/GB97/03506
/ EARLIER FILING DATE: 1997-12-22
/ EARLIER APPLICATION NUMBER: GB 9626589.7
/ EARLIER FILING DATE: 1996-12-20
/ NUMBER OF SEQ ID NOS: 18
/ SOFTWARE: PatentIn Ver. 2.1
/ SEQ ID NO 17
/ LENGTH: 19
/ TYPE: PPT
/ ORGANISM: Artificial Sequence
/ FEATURE:
/ OTHER INFORMATION: Description of Artificial Sequence: Mutant peptide
/ US-09-308-935-17

Query Match 44.6%; Score 45; DB 3; Length 19;
Best Local Similarity 90.3%; Pred. No. 1;
Matches 10; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 NVLMAMNIISK 11
Db 9 NVLMAMNIISK 19

RESULT 16
US-08-428-131-12
/ Sequence 12, Application US/08428131
/ Patent No. 5863757
/ GENERAL INFORMATION:
/ APPLICANT: La Thangue, Nicholas Barrie
/ TITLE OF INVENTION: Transcription Factor DP-1
/ NUMBER OF SEQUENCES: 14
/ CORRESPONDENCE ADDRESS:
/ ADDRESSEE: Nixon & Vanderhye
/ STREET: 1100 No. 5863757th Glebe Road, 8th Floor
/ CITY: Arlington
/ STATE: Virginia
/ COUNTRY: U.S.A.
/ ZIP: 22201-4714
/ COMPUTER READABLE FORM: disk
/ MEDIUM TYPE: Floppy
/ COMPUTER: IBM PC compatible
/ OPERATING SYSTEM: PC-DOS/MS-DOS
/ SOFTWARE: PatentIn Release #1.0, Version #1.25 (BPO)
/ CURRENT APPLICATION DATA:
/ APPLICATION NUMBER: US/08/428,131
/ FILING DATE: 23-JUN-1995
/ CLASSIFICATION: 514
```

ATTORNEY/AGENT INFORMATION:  
NAME: Arthur R. Crawford  
REGISTRATION NUMBER: 25,327  
REFERENCE/DOCKET NUMBER: 117-181  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (703) 816-4000  
TELEFAX: (703) 816-4100  
INFORMATION FOR SEQ ID NO: 12:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 73 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-428-131-12

Query Match 44.1%; Score 44.5; DB 2; Length 73;  
Best Local Similarity 42.9%; Pred. No. 5.2;  
Matches 9; Conservative 6; Mismatches 5; Indels 1; Gaps 1;

Qy 1 NVLMAMNIISKE-KKEIKWIG 20  
Db 15 NVLEGIQLIAKSKNHIQWL 35

RESULT 17  
US-09-578-596-12  
Sequence 12, Application US/09078596  
Patent No. 6150116  
GENERAL INFORMATION:  
APPLICANT: La Thangue, Nicholas Barrie  
TITLE OF INVENTION: Transcription Factor DP-1  
NUMBER OF SEQUENCES: 14  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Nixon & Vanderhye  
STREET: 1100 No. 6150116th Glebe Road, 8th Floor  
CITY: Arlington  
STATE: Virginia  
COUNTRY: U.S.A.  
ZIP: 22201-4714  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25 (EPO)  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/078,596  
FILING DATE:  
CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US/08/428,131  
FILING DATE: 23-JUN-1995  
ATTORNEY/AGENT INFORMATION:  
NAME: Arthur R. Crawford  
REGISTRATION NUMBER: 25,327  
REFERENCE/DOCKET NUMBER: 117-181  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (703) 816-4000  
TELEFAX: (703) 816-4100  
INFORMATION FOR SEQ ID NO: 12:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 73 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-09-078-596-12

Query Match 44.1%; Score 44.5; DB 3; Length 73;  
Best Local Similarity 42.9%; Pred. No. 5.2;  
Matches 9; Conservative 6; Mismatches 5; Indels 1; Gaps 1;

Qy 1 NVLMAMNIISKE-KKEIKWIG 20

ATTORNEY/AGENT INFORMATION:  
NAME: Arthur R. Crawford  
REGISTRATION NUMBER: 25,327  
REFERENCE/DOCKET NUMBER: 117-181  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (703) 816-4000  
TELEFAX: (703) 816-4100  
INFORMATION FOR SEQ ID NO: 12:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 73 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-428-131-12

Query Match 44.1%; Score 44.5; DB 2; Length 73;  
Best Local Similarity 42.9%; Pred. No. 5.2;  
Matches 9; Conservative 6; Mismatches 5; Indels 1; Gaps 1;

Qy 1 NVLMAMNIISKE-KKEIKWIG 20  
Db 15 NVLEGIQLIAKSKNHIQWL 35

RESULT 18  
US-08-894-139-5  
Sequence 5, Application US/08894139  
Patent No. 6448376  
GENERAL INFORMATION:  
APPLICANT: LA THANGUE, NICHOLAS B.  
APPLICANT: BERNARDS, RENE  
APPLICANT: HIJMAN, ELEANORE M.  
TITLE OF INVENTION: TRANSCRIPTION FACTOR E2F-5  
NUMBER OF SEQUENCES: 25  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: NIXON & VANDERHYE P.C.  
STREET: 1100 NORTH GLEBE ROAD  
CITY: ARLINGTON  
STATE: VIRGINIA  
COUNTRY: U.S.A.  
ZIP: 22201-4714  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/894,139  
FILING DATE: 13-AUG-1997  
CLASSIFICATION: 536  
ATTORNEY/AGENT INFORMATION:  
NAME: WILSON, MARY J.  
REGISTRATION NUMBER: 32,955  
REFERENCE/DOCKET NUMBER: 620-22  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (703) 816-4000  
TELEFAX: (703) 816-4100  
INFORMATION FOR SEQ ID NO: 5:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 74 amino acids  
TYPE: amino acid  
STRANDEDNESS:  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
US-08-894-139-5

Query Match 44.1%; Score 44.5; DB 4; Length 74;  
Best Local Similarity 42.9%; Pred. No. 5.2;  
Matches 9; Conservative 6; Mismatches 5; Indels 1; Gaps 1;

Qy 1 NVLMAMNIISKE-KKEIKWIG 20  
Db 53 NVLEGIQLIAKSKNHIQWL 73

RESULT 19  
US-08-894-139-7  
Sequence 7, Application US/08894139  
Patent No. 6448376  
GENERAL INFORMATION:  
APPLICANT: LA THANGUE, NICHOLAS B.  
APPLICANT: BERNARDS, RENE  
APPLICANT: HIJMAN, ELEANORE M.  
TITLE OF INVENTION: TRANSCRIPTION FACTOR E2F-5  
NUMBER OF SEQUENCES: 25  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: NIXON & VANDERHYE P.C.  
STREET: 1100 NORTH GLEBE ROAD  
CITY: ARLINGTON  
STATE: VIRGINIA  
COUNTRY: U.S.A.  
ZIP: 22201-4714  
COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/894,139  
; FILING DATE: 13-AUG-1997  
; CLASSIFICATION: 536  
; ATTORNEY/AGENT INFORMATION:  
; NAME: WILSON, MARY J.  
; REGISTRATION NUMBER: 32,955  
; REFERENCE/DOCKET NUMBER: 620-22  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (703) 816-4000  
; TELEFAX: (703) 816-4100  
; INFORMATION FOR SEQ ID NO: 7:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 74 amino acids  
; TYPE: amino acid  
; STRANDEDNESS:  
; TOPOLOGY: linear  
; MOLECULE TYPE: peptide  
US-08-894-139-7

Query Match 42.1%; Score 42.5; DB 4; Length 74;  
Best Local Similarity 38.1%; Pred. No. 11;  
Matches 8; Conservative 7; Mismatches 5; Indels 1; Gaps 1;

QY 1 NVLMAMNIIISKE-KKEIKWIG 20  
|||:::|:|:|:|:  
Db 53 NVLEGIHLIKKSKNHVQWNG 73

## RESULT 20

US-09-308-935-2  
; Sequence 2, Application US/09308935  
; Patent No. 6268334  
; GENERAL INFORMATION:  
; APPLICANT: La Thangue, Nicholas B  
; APPLICANT: Bandata, Iaseantha R  
; TITLE OF INVENTION: Peptide antagonists of DP transcription factors  
; FILE REFERENCE: 620-67  
; CURRENT APPLICATION NUMBER: US/09/308,935  
; CURRENT FILING DATE: 1999-05-27  
; EARLIER APPLICATION NUMBER: PCT/GB97/03506  
; EARLIER FILING DATE: 1997-12-22  
; EARLIER APPLICATION NUMBER: GB 9626589.7  
; EARLIER FILING DATE: 1996-12-20  
; NUMBER OF SEQ ID NOS: 18  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 2  
; LENGTH: 9  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Description of Artificial Sequence: Synthetic  
US-09-308-935-2

Query Match 41.6%; Score 42; DB 3; Length 9;  
Best Local Similarity 100.0%; Pred. No. 2.5e+05;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 NVLMAMNII 9  
|||||||  
Db 1 NVLMAMNII 9

## RESULT 21

US-08-894-139-9  
; Sequence 9, Application US/08894139  
; Patent No. 6448376  
; GENERAL INFORMATION:  
; APPLICANT: LA THANGUE, NICHOLAS B.

; APPLICANT: BERNARDS, RENE  
; APPLICANT: HIJMANS, ELEANORE M.  
; TITLE OF INVENTION: TRANSCRIPTION FACTOR E2F-5  
; NUMBER OF SEQUENCES: 25  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: NIXON & VANDERHUYE P.C.  
; STREET: 1100 NORTH GLEBE ROAD  
; CITY: ARLINGTON  
; STATE: VIRGINIA  
; COUNTRY: U.S.A.  
; ZIP: 22201-4714  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/894,139  
; FILING DATE: 13-AUG-1997  
; CLASSIFICATION: 536  
; ATTORNEY/AGENT INFORMATION:  
; NAME: WILSON, MARY J.  
; REGISTRATION NUMBER: 32,955  
; REFERENCE/DOCKET NUMBER: 620-22  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (703) 816-4000  
; TELEFAX: (703) 816-4100  
; INFORMATION FOR SEQ ID NO: 9:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 75 amino acids  
; TYPE: amino acid  
; STRANDEDNESS:  
; TOPOLOGY: linear  
; MOLECULE TYPE: peptide  
US-08-894-139-9

Query Match 39.1%; Score 39.5; DB 4; Length 75;  
Best Local Similarity 42.9%; Pred. No. 32;  
Matches 9; Conservative 5; Mismatches 6; Indels 1; Gaps 1;

QY 1 NVLMAMNIIISKE-KKEIKWIG 20  
|||:::|:|:|:|:  
Db 54 NVLEGIHLIKKSKNSIQWKG 74

## RESULT 22

US-09-328-352-7148  
; Sequence 7148, Application US/09328352  
; Patent No. 6562958  
; GENERAL INFORMATION:  
; APPLICANT: Gary L. Breton et al.  
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO ACINETOBACTER  
; FILE REFERENCE: GTC99-03PA  
; CURRENT APPLICATION NUMBER: US/09/328,352  
; CURRENT FILING DATE: 1999-06-04  
; NUMBER OF SEQ ID NOS: 8252  
; SEQ ID NO 7148  
; LENGTH: 78  
; TYPE: PRT  
; ORGANISM: Acinetobacter baumannii  
US-09-328-352-7148

Query Match 38.6%; Score 39; DB 4; Length 78;  
Best Local Similarity 31.2%; Pred. No. 39;  
Matches 5; Conservative 8; Mismatches 3; Indels 0; Gaps 0;

QY 4 MAMNIIISKEKKEIKWI 19  
:|:|:|:|:|:|:|:  
Db 40 VAMSLITRPRKALDWL 55

## RESULT 23

US-09-107-532A-5431  
; Sequence 5431, Application US/09107532A  
; Patent No. 6583275  
; GENERAL INFORMATION:  
; APPLICANT: Lynn A Doucette-Stamm and David Bush  
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO  
; ENTEROCOCCUS FAECIUM FOR DIAGNOSTICS AND THERAPEUTICS  
; NUMBER OF SEQUENCES: 7310  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: GENOME THERAPEUTICS CORPORATION  
; STREET: 100 Beaver Street  
; CITY: Waltham  
; STATE: Massachusetts  
; COUNTRY: USA  
; ZIP: 02354  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: CD-ROM ISO9660  
; COMPUTER: PC  
; OPERATING SYSTEM: <unknown>  
; SOFTWARE: ASCII  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/107,532A  
; FILING DATE: 30-Jun-1998  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 60/085,598  
; FILING DATE: 14 May 1998  
; APPLICATION NUMBER: 60/051571  
; FILING DATE: July 2, 1997  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Ariniello, Pamela Deneke  
; REGISTRATION NUMBER: 40,489  
; REFERENCE/DOCKET NUMBER: GTC-012  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (781)893-5007  
; TELEFAX: (781)893-8277  
; INFORMATION FOR SEQ ID NO: 5431:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 87 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
; HYPOTHETICAL: YES  
; ORIGINAL SOURCE:  
; ORGANISM: Enterococcus faecium  
; FEATURE:  
; NAME/KEY: misc feature  
; LOCATION: (B) LOCATION 1...87  
; SEQUENCE DESCRIPTION: SEQ ID NO: 5431:  
US-09-107-532A-5431  
Query Match 36 68; Score 37; DB 4; Length 87;  
Best Local Similarity 46.7%; Pred. No. 90;  
Matches 7; Conservative 5; Mismatches 3; Indels 0; Gaps 0;  
Qy 3 LMANNIISKEKKEIK 17  
|.:|:|:|:|:|  
Db 23 LLAIQELSKKKLK 37  
RESULT 24  
US-07-948-357-3  
; Sequence 3, Application US/07948357  
; Patent No. 5547932  
; GENERAL INFORMATION:  
; APPLICANT: Curiel, David T.  
; APPLICANT: Birnstiel, Max L.  
; APPLICANT: Cotten, Matthew  
; APPLICANT: Wagner, Ernst  
; APPLICANT: Zatloukal, Kurt  
; APPLICANT: Plank, Christian  
; APPLICANT: Oberhauser, Berndt  
; APPLICANT: Schmidt, Walter G.M.  
; TITLE OF INVENTION: Composition for Introducing Nucleic Acid  
; COMPLEXES INTO HIGHER EUKARYOTIC CELLS  
; NUMBER OF SEQUENCES: 13  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Sterne, Kessler, Goldstein & Fox  
; STREET: 1225 Connecticut Avenue  
; CITY: Washington  
; STATE: D.C.  
; COUNTRY: U.S.A.  
; ZIP: 20036  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent in Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/07/948,357  
; FILING DATE: 19920923  
; CLASSIFICATION: 514  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Esmond, Robert W.  
; REGISTRATION NUMBER: 32,893  
; REFERENCE/DOCKET NUMBER: 0652.0940004  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (202) 466-0800  
; TELEFAX: (202) 833-8716  
; INFORMATION FOR SEQ ID NO: 3:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 26 amino acids  
; TYPE: AMINO ACID  
; STRANDEDNESS: single  
; TOPOLOGY: both  
; MOLECULE TYPE: peptide  
US-07-948-357-3  
Query Match 34.7%; Score 35; DB 1; Length 26;  
Best Local Similarity 50.0%; Pred. No. 50;  
Matches 8; Conservative 2; Mismatches 6; Indels 0; Gaps 0;  
Qy 4 MAMNIISKEKKEIKWI 19  
|:|:|:|:|:|  
Db 1 MAQDIISTIGDLVKWI 16

US-09-107-532A-5431  
; Sequence 5431, Application US/09107532A  
; Patent No. 6583275  
; GENERAL INFORMATION:  
; APPLICANT: Lynn A Doucette-Stamm and David Bush  
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO  
; ENTEROCOCCUS FAECIUM FOR DIAGNOSTICS AND THERAPEUTICS  
; NUMBER OF SEQUENCES: 7310  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: GENOME THERAPEUTICS CORPORATION  
; STREET: 100 Beaver Street  
; CITY: Waltham  
; STATE: Massachusetts  
; COUNTRY: USA  
; ZIP: 02354  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: CD-ROM ISO9660  
; COMPUTER: PC  
; OPERATING SYSTEM: <unknown>  
; SOFTWARE: ASCII  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/107,532A  
; FILING DATE: 30-Jun-1998  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 60/085,598  
; FILING DATE: 14 May 1998  
; APPLICATION NUMBER: 60/051571  
; FILING DATE: July 2, 1997  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Ariniello, Pamela Deneke  
; REGISTRATION NUMBER: 40,489  
; REFERENCE/DOCKET NUMBER: GTC-012  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (781)893-5007  
; TELEFAX: (781)893-8277  
; INFORMATION FOR SEQ ID NO: 5431:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 87 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
; HYPOTHETICAL: YES  
; ORIGINAL SOURCE:  
; ORGANISM: Enterococcus faecium  
; FEATURE:  
; NAME/KEY: misc feature  
; LOCATION: (B) LOCATION 1...87  
; SEQUENCE DESCRIPTION: SEQ ID NO: 5431:  
US-09-107-532A-5431  
Query Match 36 68; Score 37; DB 4; Length 87;  
Best Local Similarity 46.7%; Pred. No. 90;  
Matches 7; Conservative 5; Mismatches 3; Indels 0; Gaps 0;  
Qy 3 LMANNIISKEKKEIK 17  
|.:|:|:|:|:|  
Db 23 LLAIQELSKKKLK 37  
RESULT 24  
US-07-948-357-3  
; Sequence 3, Application US/07948357  
; Patent No. 5547932  
; GENERAL INFORMATION:  
; APPLICANT: Curiel, David T.  
; APPLICANT: Birnstiel, Max L.  
; APPLICANT: Cotten, Matthew  
; APPLICANT: Wagner, Ernst  
; APPLICANT: Zatloukal, Kurt  
; APPLICANT: Plank, Christian  
; APPLICANT: Oberhauser, Berndt  
; APPLICANT: Schmidt, Walter G.M.  
; TITLE OF INVENTION: Composition for Introducing Nucleic Acid

```
/
/ FILING DATE:
/ CLASSIFICATION: 435
/ PRIOR APPLICATION DATA:
/ APPLICATION NUMBER: US/07/948,357
/ FILING DATE: 19920923
/ ATTORNEY/AGENT INFORMATION:
/ NAME: Esmond, Robert W.
/ REGISTRATION NUMBER: 32,893
/ REFERENCE/DOCKET NUMBER: 0652.0940004
/ TELECOMMUNICATION INFORMATION:
/ TELEPHONE: (202) 466-0800
/ TELEFAX: (202) 833-8716
/ INFORMATION FOR SEQ ID NO: 3:
/ SEQUENCE CHARACTERISTICS:
/ LENGTH: 26 amino acids
/ TYPE: amino acid
/ STRANDEDNESS: single
/ TOPOLOGY: both
/ MOLECULE TYPE: peptide
/ US-08-450-417-3

Query Match          34.7%  Score 35; DB 2; Length 26;
Best Local Similarity 50.0%; Pred. No. 50;
Matches 8; Conservative 2; Mismatches 6; Indels 0; Gaps 0;

QY      4 MAMNIISKEKKEIKWI 19
Db      1 MAQDIISTIGDLVKWI 16

RESULT 26
US-08-449-741-3
/ Sequence 3, Application US/08449741
/ Patent No. 6022735
/ GENERAL INFORMATION:
/ APPLICANT: Curiel, David T.
/ APPLICANT: Birnstiel, Max L.
/ APPLICANT: Cotten, Matthew
/ APPLICANT: Wagner, Ernst
/ APPLICANT: Zatloukal, Kurt
/ APPLICANT: Plank, Christian
/ APPLICANT: Oberhauser, Berndt
/ APPLICANT: Schmidt, Walter G.M.
/ TITLE OF INVENTION: Composition for Introducing Nucleic Acid
/ TITLE OF INVENTION: Complexes Into Higher Eucaryotic Cells
/ NUMBER OF SEQUENCES: 13
/ CORRESPONDENCE ADDRESS:
/ ADDRESSEE: Sterne, Kessler, Goldstein & Fox
/ STREET: 1100 New York Ave., Suite 600
/ CITY: Washington
/ STATE: D.C.
/ COUNTRY: U.S.A.
/ ZIP: 20005
/ COMPUTER READABLE FORM:
/ MEDIUM TYPE: Floppy disk
/ COMPUTER: IBM PC compatible
/ OPERATING SYSTEM: PC-DOS/MS-DOS
/ SOFTWARE: PatentIn Release #1.0, Version #1.25
/ CURRENT APPLICATION DATA:
/ APPLICATION NUMBER: US/08/449,741
/ FILING DATE: 25-MAY-1995
/ CLASSIFICATION: 514
/ ATTORNEY/AGENT INFORMATION:
/ NAME: Fleisher, Raz E.
/ REGISTRATION NUMBER: 34,331
/ REFERENCE/DOCKET NUMBER: 0652.0940007
/ TELECOMMUNICATION INFORMATION:
/ TELEPHONE: (202) 371-2600
/ TELEFAX: (202) 371-2540
/ INFORMATION FOR SEQ ID NO: 3:
/ SEQUENCE CHARACTERISTICS:
/ LENGTH: 26 amino acids
/ TYPE: amino acid
```

```
/
/ STRANDEDNESS: single
/ TOPOLOGY: both
/ MOLECULE TYPE: peptide
/ US-08-449-741-3

Query Match          34.7%  Score 35; DB 3; Length 26;
Best Local Similarity 50.0%; Pred. No. 50;
Matches 8; Conservative 2; Mismatches 6; Indels 0; Gaps 0;

QY      4 MAMNIISKEKKEIKWI 19
Db      1 MAQDIISTIGDLVKWI 16

RESULT 27
US-08-782-997A-5
/ Sequence 5, Application US/08782997A
/ Patent No. 6030602
/ GENERAL INFORMATION:
/ APPLICANT: Legendre, Jean-Yves
/ APPLICANT: Supersaxo, Andreas
/ APPLICANT: Trzeciak, Arnold
/ TITLE OF INVENTION: Peptide Conjugates for Transfecting
/ TITLE OF INVENTION: Cells
/ NUMBER OF SEQUENCES: 37
/ CORRESPONDENCE ADDRESS:
/ ADDRESSEE: Hoffmann-La Roche Inc.
/ STREET: 340 Kingsland Street
/ CITY: Nutley
/ STATE: New Jersey
/ COUNTRY: U.S.A.
/ ZIP: 07110
/ COMPUTER READABLE FORM:
/ MEDIUM TYPE: Floppy disk
/ COMPUTER: IBM PC compatible
/ OPERATING SYSTEM: PC-DOS/MS-DOS
/ SOFTWARE: PatentIn Release #1.0, Version #1.25
/ CURRENT APPLICATION DATA:
/ APPLICATION NUMBER: US/08/782,997A
/ FILING DATE: 14-JAN-1997
/ CLASSIFICATION: 435
/ PRIOR APPLICATION DATA:
/ APPLICATION NUMBER: EP 96100603.8
/ FILING DATE: 17-JAN-1996
/ ATTORNEY/AGENT INFORMATION:
/ NAME: Kreisler, Lewis J.
/ REGISTRATION NUMBER: 38,522
/ REFERENCE/DOCKET NUMBER: RAN 4600/73
/ TELECOMMUNICATION INFORMATION:
/ TELEPHONE: (973) 235-4387
/ TELEFAX: (973) 235-2363
/ INFORMATION FOR SEQ ID NO: 5:
/ SEQUENCE CHARACTERISTICS:
/ LENGTH: 26 amino acids
/ TYPE: amino acid
/ TOPOLOGY: linear
/ MOLECULE TYPE: peptide
/ HYPOTHETICAL: NO
/ FEATURE:
/ NAME/KEY: Modified-site
/ LOCATION: 1..2
/ OTHER INFORMATION: /note= "Position 1 is fMet."
/ US-08-782-997A-5

Query Match          34.7%  Score 35; DB 3; Length 26;
Best Local Similarity 50.0%; Pred. No. 50;
Matches 8; Conservative 2; Mismatches 6; Indels 0; Gaps 0;

QY      4 MAMNIISKEKKEIKWI 19
Db      1 MAQDIISTIGDLVKWI 16
```

RESULT 28  
US-08-449-754-3  
Sequence 3, Application US/08449754  
Patent No. 6077663  
GENERAL INFORMATION:  
APPLICANT: Curiel, David T.  
APPLICANT: Birnstiel, Max L.  
APPLICANT: Cotten, Matthew  
APPLICANT: Wagner, Ernst  
APPLICANT: Zatloukal, Kurt  
APPLICANT: Plank, Christian  
APPLICANT: Oberhauser, Berndt  
APPLICANT: Schmidt, Walter G.M.  
TITLE OF INVENTION: Composition for Introducing Nucleic Acid  
Complexes Into Higher Eucaryotic Cells  
NUMBER OF SEQUENCES: 13  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Sterne, Kessler, Goldstein & Fox P.L.L.C.  
STREET: 1100 New York Ave., N.W., Suite 600  
CITY: Washington  
STATE: D.C.  
COUNTRY: U.S.A.  
ZIP: 20005-3934  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/465,646  
FILING DATE: HEREWITH  
CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/449,754  
FILING DATE: May 25, 1995  
CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 07/937,788  
FILING DATE: September 2, 1992  
CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 07/864,759  
FILING DATE: April 7, 1992  
CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 07/827,102  
FILING DATE: January 30, 1992  
CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 07/827,103  
FILING DATE: January 30, 1992  
CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 07/767,788  
FILING DATE: September 30, 1991  
CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 07/768,039  
FILING DATE: September 30, 1991  
CLASSIFICATION:  
ATTORNEY/AGENT INFORMATION:  
NAME: Kamage, Andrea J.  
REGISTRATION NUMBER: 43,703  
REFERENCE/DOCKET NUMBER: 0652.0940009  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (202) 371-2600  
TELEFAX: (202) 371-6566  
INFORMATION FOR SEQ ID NO: 3:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 26 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: both  
MOLECULE TYPE: peptide  
US-09-465-646-3

Query Match 34.7%; Score 35; DB 3; Length 26;  
Best Local Similarity 50.0%; Pred. No. 50;  
Matches 8; Conservative 2; Mismatches 6; Indels 0; Gaps 0;

Qy 4 MAMNIISSKEKEIKWI 19  
Db 1 MAQDIISTIGDLVKWI 16

RESULT 30

US-08-449-754-3  
Sequence 3, Application US/08449754  
Patent No. 6077663  
GENERAL INFORMATION:  
APPLICANT: Curiel, David T.  
APPLICANT: Birnstiel, Max L.  
APPLICANT: Cotten, Matthew  
APPLICANT: Wagner, Ernst  
APPLICANT: Zatloukal, Kurt  
APPLICANT: Plank, Christian  
APPLICANT: Oberhauser, Berndt  
APPLICANT: Schmidt, Walter G.M.  
TITLE OF INVENTION: Composition for Introducing Nucleic Acid  
Complexes Into Higher Eucaryotic Cells  
NUMBER OF SEQUENCES: 13  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Sterne, Kessler, Goldstein & Fox  
STREET: 1225 Connecticut Avenue  
CITY: Washington  
STATE: D.C.  
COUNTRY: U.S.A.  
ZIP: 20036  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/449,754  
FILING DATE:  
CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US/07/948,357  
FILING DATE: 1992-09-23  
ATTORNEY/AGENT INFORMATION:  
NAME: Esmond, Robert W.  
REGISTRATION NUMBER: 32,893  
REFERENCE/DOCKET NUMBER: 0652.0940004  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (202) 466-0800  
TELEFAX: (202) 833-8716  
INFORMATION FOR SEQ ID NO: 3:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 26 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: both  
MOLECULE TYPE: peptide  
US-08-449-754-3

Query Match 34.7%; Score 35; DB 3; Length 26;  
Best Local Similarity 50.0%; Pred. No. 50;  
Matches 8; Conservative 2; Mismatches 6; Indels 0; Gaps 0;

Qy 4 MAMNIISSKEKEIKWI 19  
Db 1 MAQDIISTIGDLVKWI 16

RESULT 29  
US-09-465-646-3  
Sequence 3, Application US/09465646  
Patent No. 6274322  
GENERAL INFORMATION:  
APPLICANT: Curiel, David T.  
APPLICANT: Birnstiel, Max L.  
APPLICANT: Cotten, Matthew  
APPLICANT: Wagner, Ernst  
APPLICANT: Zatloukal, Kurt  
APPLICANT: Plank, Christian  
APPLICANT: Oberhauser, Berndt  
APPLICANT: Schmidt, Walter G.M.

```
US-07-948-357-4
; Sequence 4, Application US/07948357
; Patent No. 5547932
; GENERAL INFORMATION:
; APPLICANT: Curiel, David T.
; APPLICANT: Birnstiel, Max L.
; APPLICANT: Cotten, Matthew
; APPLICANT: Wagner, Ernst
; APPLICANT: Zatloukal, Kurt
; APPLICANT: Plank, Christian
; APPLICANT: Oberhauser, Berndt
; APPLICANT: Schmidt, Walter G.M.
; TITLE OF INVENTION: Composition for Introducing Nucleic Acid
; TITLE OF INVENTION: Complexes Into Higher Eucaryotic Cells
; NUMBER OF SEQUENCES: 13
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Sterne, Kessler, Goldstein & Fox
; STREET: 1225 Connecticut Avenue
; CITY: Washington
; STATE: D.C.
; COUNTRY: U.S.A.
; ZIP: 20036
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/948,357
; FILING DATE: 19920923
; CLASSIFICATION: 514
; ATTORNEY/AGENT INFORMATION:
; NAME: Esmond, Robert W.
; REGISTRATION NUMBER: 32,893
; REFERENCE/DOCKET NUMBER: 0652.0940004
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 466-0800
; TELEFAX: (202) 833-8716
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 36 amino acids
; TYPE: AMINO ACID
; STRANDEDNESS: single
; TOPOLOGY: both
; MOLECULE TYPE: peptide
; US-07-948-357-4

Query Match 34.7%; Score 35; DB 1; Length 36;
Best Local Similarity 50.0%; Pred. No. 71;
Matches 8; Conservative 2; Mismatches 6; Indels 0; Gaps 0;

Qy 4 MAMNIISKEKKEIKWI 19
Db 1 MAQDIISTIGDLVKWI 16

RESULT 31
US-08-450-417-4
; Sequence 4, Application US/08450417
; Patent No. 5981273
; GENERAL INFORMATION:
; APPLICANT: Curiel, David T.
; APPLICANT: Birnstiel, Max L.
; APPLICANT: Cotten, Matthew
; APPLICANT: Wagner, Ernst
; APPLICANT: Zatloukal, Kurt
; APPLICANT: Plank, Christian
; APPLICANT: Oberhauser, Berndt
; APPLICANT: Schmidt, Walter G.M.
; TITLE OF INVENTION: Composition for Introducing Nucleic Acid
; TITLE OF INVENTION: Complexes Into Higher Eucaryotic Cells
; NUMBER OF SEQUENCES: 13
; CORRESPONDENCE ADDRESS:
```

```
ADDRESSEE: Sterne, Kessler, Goldstein & Fox
STREET: 1225 Connecticut Avenue
CITY: Washington
STATE: D.C.
COUNTRY: U.S.A.
ZIP: 20036
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/450,417
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/07/948,357
FILING DATE: 19920923
ATTORNEY/AGENT INFORMATION:
NAME: Esmond, Robert W.
REGISTRATION NUMBER: 32,893
REFERENCE/DOCKET NUMBER: 0652.0940004
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202) 466-0800
TELEFAX: (202) 833-8716
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 36 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: both
MOLECULE TYPE: peptide
US-08-450-417-4

Query Match 34.7%; Score 35; DB 2; Length 36;
Best Local Similarity 50.0%; Pred. No. 71;
Matches 8; Conservative 2; Mismatches 6; Indels 0; Gaps 0;

Qy 4 MAMNIISKEKKEIKWI 19
Db 1 MAQDIISTIGDLVKWI 16

RESULT 32
US-08-449-741-4
; Sequence 4, Application US/08449741
; Patent No. 6022735
; GENERAL INFORMATION:
; APPLICANT: Curiel, David T.
; APPLICANT: Birnstiel, Max L.
; APPLICANT: Cotten, Matthew
; APPLICANT: Wagner, Ernst
; APPLICANT: Zatloukal, Kurt
; APPLICANT: Plank, Christian
; APPLICANT: Oberhauser, Berndt
; APPLICANT: Schmidt, Walter G.M.
; TITLE OF INVENTION: Composition for Introducing Nucleic Acid
; TITLE OF INVENTION: Complexes Into Higher Eucaryotic Cells
; NUMBER OF SEQUENCES: 13
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Sterne, Kessler, Goldstein & Fox
; STREET: 1100 New York Ave., Suite 600
; CITY: Washington
; STATE: D.C.
; COUNTRY: U.S.A.
; ZIP: 20005
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/449,741
```





```
;
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Kamege, Andrea J.
; REGISTRATION NUMBER: 43,703
; REFERENCE/DOCKET NUMBER: 0652.0940009
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 371-2600
; TELEFAX: (202) 371-6566
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 36 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: both
; MOLECULE TYPE: peptide
; US-09-465-646-4
;
; Query Match 34.7%; Score 35; DB 3; Length 36;
; Best Local Similarity 50.0%; Pred. No. 71;
; Matches 8; Conservative 2; Mismatches 6; Indels 0; Gaps 0;
;
; QY 4 MAMNIISKEKEIKWI 19
; ||:||||:
; Db 1 MAQDIISTIGDLVKWI 16
;
; RESULT 35
; US-08-194-338-14
; Sequence 14, Application US/08194338
; Patent No. 5474898
; GENERAL INFORMATION:
; APPLICANT: Venter, John C.
; APPLICANT: Fraser, Claire M.
; APPLICANT: McCombie, William R.
; TITLE OF INVENTION: OCTOPAMINE RECEPTOR
; NUMBER OF SEQUENCES: 16
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Knobbe, Martens, Olson and Bear
; STREET: 620 Newport Center Drive, Sixteenth Floor
; CITY: Newport Beach
; STATE: CA
; COUNTRY: USA
; ZIP: 92660
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/194,338
; FILING DATE: 08-FEB-1994
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/676,174
; FILING DATE: 28-MAR-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: Israelson, Ned A.
; REGISTRATION NUMBER: 29,655
; REFERENCE/DOCKET NUMBER: NIH01.001DV1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (619) 235-8550
; TELEFAX: (619) 235-0176
; INFORMATION FOR SEQ ID NO: 14:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 63 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; FRAGMENT TYPE: internal
; US-08-194-338-14
;
; Query Match 34.7%; Score 35; DB 1; Length 63;
; Best Local Similarity 50.0%; Pred. No. 1.3e+02;
; Matches 6; Conservative 5; Mismatches 1; Indels 0; Gaps 0;
;
; QY 1 NVLMAMNIISKE 12
; |||:|:|:|:|:
; Db 3 NVLVALNLAAD 14
;
; RESULT 36
; US-09-308-935-11
; Sequence 11, Application US/09308935
; Patent No. 6268334
; GENERAL INFORMATION:
; APPLICANT: La Thangue, Nicholas B
; APPLICANT: Bandara, Lasantha R
; TITLE OF INVENTION: Peptide antagonists of DP transcription factors
; FILE REFERENCE: 620-67
; CURRENT APPLICATION NUMBER: US/09/308,935
; CURRENT FILING DATE: 1999-05-27
; EARLIER APPLICATION NUMBER: PCT/GB97/03506
; EARLIER FILING DATE: 1997-12-22
; EARLIER APPLICATION NUMBER: GB 9626589.7
; EARLIER FILING DATE: 1996-12-20
; NUMBER OF SEQ ID NOS: 18
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 11
; LENGTH: 14
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Synthetic
; US-09-308-935-11
;
; Query Match 33.7%; Score 34; DB 3; Length 14;
; Best Local Similarity 100.0%; Pred. No. 37;
; Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
;
; QY 1 NVLMAMN 7
; |||||
; Db 8 NVLMAMN 14
;
; RESULT 37
; US-09-230-548-15
; Sequence 15, Application US/09230548
; Patent No. 6326466
; GENERAL INFORMATION:
; APPLICANT: Bottaro, Donald P.
; APPLICANT: Petryshyn, Raymond
; APPLICANT: The Government of the United States of America
; APPLICANT: as represented by The Secretary,
; APPLICANT: Department of Health and Human Services
; TITLE OF INVENTION: Double-Stranded RNA Dependent Protein Kinase Derived
; TITLE OF INVENTION: Peptides to Promote Proliferation of Cells and Tissues
; TITLE OF INVENTION: in a Controlled Manner
; FILE REFERENCE: 015280-286200US
; CURRENT APPLICATION NUMBER: US/09/230,548
; CURRENT FILING DATE: 1999-07-23
; EARLIER APPLICATION NUMBER: US 60/023,307
; EARLIER FILING DATE: 1996-07-30
; EARLIER APPLICATION NUMBER: WO PCT/US97/14350
; EARLIER FILING DATE: 1997-07-29
; NUMBER OF SEQ ID NOS: 31
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 15
; LENGTH: 21
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence:double-stranded
; OTHER INFORMATION: RNA dependent protein kinase (PKR) peptide
; US-09-230-548-15
```

```
; OTHER INFORMATION: antagonist, human PKR site 1 (hR1 peptide)
; US-09-230-548-15
Query Match      33.7%; Score 34; DB 4; Length 21;
Best Local Similarity 54.5%; Pred. No. 57;
Matches 6; Conservative 4; Mismatches 1; Indels 0; Gaps 0;

QY 4 MAMNIISKEKK 14
   :|:|:|:|:|
Db 11 LAVEILNKEKK 21

RESULT 38
US-09-230-548-26
; Sequence 26, Application US/09230548
; Patent No. 6326466
; GENERAL INFORMATION:
; APPLICANT: Bottaro, Donald P.
; APPLICANT: Petryshyn, Raymond
; APPLICANT: The Government of the United States of America
; APPLICANT: as represented by The Secretary,
; APPLICANT: Department of Health and Human Services
; TITLE OF INVENTION: Double-Stranded RNA Dependent Protein Kinase Derived
; TITLE OF INVENTION: Peptides to Promote Proliferation of Cells and Tissues
; TITLE OF INVENTION: in a Controlled Manner
; FILE REFERENCE: 015280-286200US
; CURRENT APPLICATION NUMBER: US/09/230,548
; CURRENT FILING DATE: 1999-07-23
; EARLIER APPLICATION NUMBER: US 60/023,307
; EARLIER FILING DATE: 1996-07-30
; EARLIER APPLICATION NUMBER: WO PCT/US97/14350
; EARLIER FILING DATE: 1997-07-29
; NUMBER OF SEQ ID NOS: 31
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 26
; LENGTH: 26
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence:double-stranded
; OTHER INFORMATION: RNA dependent protein kinase (PKR) peptide
; OTHER INFORMATION: antagonist, human PKR site 1
; US-09-230-548-26

Query Match      33.7%; Score 34; DB 4; Length 26;
Best Local Similarity 54.5%; Pred. No. 72;
Matches 6; Conservative 4; Mismatches 1; Indels 0; Gaps 0;

QY 4 MAMNIISKEKK 14
   :|:|:|:|:|
Db 16 LAVEILNKEKK 26

RESULT 39
US-09-230-548-31
; Sequence 31, Application US/09230548
; Patent No. 6326466
; GENERAL INFORMATION:
; APPLICANT: Bottaro, Donald P.
; APPLICANT: Petryshyn, Raymond
; APPLICANT: The Government of the United States of America
; APPLICANT: as represented by The Secretary,
; APPLICANT: Department of Health and Human Services
; TITLE OF INVENTION: Double-Stranded RNA Dependent Protein Kinase Derived
; TITLE OF INVENTION: Peptides to Promote Proliferation of Cells and Tissues
; TITLE OF INVENTION: in a Controlled Manner
; FILE REFERENCE: 015280-286200US
; CURRENT APPLICATION NUMBER: US/09/230,548
; CURRENT FILING DATE: 1999-07-23
; EARLIER APPLICATION NUMBER: US 60/023,307
; EARLIER FILING DATE: 1996-07-30
; EARLIER APPLICATION NUMBER: WO PCT/US97/14350
; EARLIER FILING DATE: 1997-07-29
; NUMBER OF SEQ ID NOS: 31
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 31
; LENGTH: 31
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence:double-stranded
; OTHER INFORMATION: RNA dependent protein kinase (PKR) peptide
; OTHER INFORMATION: antagonist, cell-permeable PKR peptide antagonist
; OTHER INFORMATION: subsequences
; US-09-230-548-31

Query Match      33.7%; Score 34; DB 4; Length 46;
Best Local Similarity 54.5%; Pred. No. 1.3e+02;
Matches 6; Conservative 4; Mismatches 1; Indels 0; Gaps 0;

QY 4 MAMNIISKEKK 14
   :|:|:|:|:|
Db 36 LAVEILNKEKK 46

RESULT 41
US-08-894-139-8
; Sequence 8, Application US/08894139
; Patent No. 6448376
; GENERAL INFORMATION:
; APPLICANT: LA THANGUE, NICHOLAS B.
; APPLICANT: BERNARDS, RENE
; APPLICANT: HJUMANS, ELEANORE M.
; TITLE OF INVENTION: TRANSCRIPTION FACTOR E2F-5
```

```

;
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/343,443B
; FILING DATE: 18-NOV-1994
; CLASSIFICATION: 514
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/FR93/00494
; FILING DATE: 19-MAY-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: FR 92/06123
; FILING DATE: 20-MAY-1992
; ATTORNEY/AGENT INFORMATION:
; NAME: Weiser, Gerard J.
; REGISTRATION NUMBER: 19,763
; REFERENCE/DOCKET NUMBER: 989,6121P
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 215-875-8383
; TELEFAX: 215-875-8394
; INFORMATION FOR SEQ ID NO: 8:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 86 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
;
; US-08-343-443B-8
;
; Query Match 33.2%; Score 33.5; DB 2; Length 86;
; Best Local Similarity 38.9%; Pred. No. 3.1e+02;
; Matches 7; Conservative 5; Mismatches 3; Indels 3; Caps 1;
;
; QY 5 AMNIIISKEKEIK--WI 19
; | : | | : | : | : | : |
; DB 21 AIGIIKKDKRTMKPKIWL 38
;
; RESULT 43
; US-08-743-975-9
; Sequence 9, Application US/08743975
; Patent No. 6057434
; GENERAL INFORMATION:
; APPLICANT: Jian Ni, Guo-Liang Yu and Reiner Gentz
; TITLE OF INVENTION: Mammary Transforming Protein
; NUMBER OF SEQUENCES: 9
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: CARELLA, BYRNE, BAIN, GILFILLAN,
; ADDRESSEE: CECCHI, STEWART & OLSTEIN
; STREET: 6 BECKER FARM ROAD
; CITY: ROSELAND
; STATE: NEW JERSEY
; COUNTRY: USA
; ZIP: 07068
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5 INCH DISKETTE
; COMPUTER: IBM PS/2
; OPERATING SYSTEM: MS-DOS
; SOFTWARE: WORD PERFECT 5.1
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/743,975
; FILING DATE: 01 NOVEMBER 1996
; CLASSIFICATION: 514
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 60/006,187
; FILING DATE: 02 NOVEMBER 1995
; ATTORNEY/AGENT INFORMATION:
; NAME: MULLINS, J.G.
; REGISTRATION NUMBER: 33,073
; REFERENCE/DOCKET NUMBER: 325800-507 (PF212)
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 201-994-1700
; TELEFAX: 201-994-1744
; INFORMATION FOR SEQ ID NO: 9:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 61 AMINO ACIDS
; TYPE: AMINO ACID
; STRANDEDNESS:

```

```

; TOPOLOGY: LINEAR
; MOLECULE TYPE: PROTEIN
US-08-743-975-9
Query Match 32.7%; Score 33; DB 3; Length 61;
Best Local Similarity 31.6%; Pred. No. 2.6e+02;
Matches 6; Conservative 8; Mismatches 3; Indels 2; Gaps 1;

Qy 2 VLMAMNIISKEKEIKWIG 20
| : : : : : : : : : :
Db 14 VVFSINLLSRPERE--WEG 30

RESULT 45
US-09-263-811-9
; Sequence 9, Application US/09263811
; Patent No. 6482922
; GENERAL INFORMATION:
; APPLICANT: Jian Ni et al.
; TITLE OF INVENTION: Mammary Transforming Protein
; NUMBER OF SEQUENCES: 9
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: HUMAN GENOME SCIENCES, INC.
; STREET: 9410 KEY WEST AVENUE
; CITY: ROCKVILLE
; STATE: MARYLAND
; COUNTRY: USA
; ZIP: 20850
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5 INCH DISKETTE
; COMPUTER: IBM PS/2
; OPERATING SYSTEM: MS-DOS
; SOFTWARE: WORD PERFECT 5.1
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/263,811
; FILING DATE: MAR-08-1999
; CLASSIFICATION: 536
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/743,975
; FILING DATE: 01-NOV-1996
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 60/006,187
; FILING DATE: 02 NOVEMBER 1995
; ATTORNEY/AGENT INFORMATION:
; NAME: MICHELE M. WALES
; REGISTRATION NUMBER: 43,975
; REFERENCE/DOCKET NUMBER: PF212D1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 301-309-8504
; TELEFAX: 301-309-8439
; INFORMATION FOR SEQ ID NO: 9:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 61 AMINO ACIDS
; TYPE: AMINO ACID
; STRANDEDNESS:
; TOPOLOGY: LINEAR
; MOLECULE TYPE: PROTEIN
US-09-263-811-9
Query Match 32.7%; Score 33; DB 4; Length 61;
Best Local Similarity 31.6%; Pred. No. 2.6e+02;
Matches 6; Conservative 8; Mismatches 3; Indels 2; Gaps 1;

Qy 2 VLMAMNIISKEKEIKWIG 20
| : : : : : : : : : :
Db 14 VVFSINLLSRPERE--WEG 30

RESULT 45
US-09-205-258-381
; Sequence 381, Application US/09205258
; Patent No. 6525174
; GENERAL INFORMATION:
; APPLICANT: Young et al.
; TITLE OF INVENTION: 207 Human Secreted Proteins
; FILE REFERENCE: PZ007P1
; CURRENT APPLICATION NUMBER: US/09/205,258
; CURRENT FILING DATE: 1998-12-04
; EARLIER APPLICATION NUMBER: PCT/US98/11422
; EARLIER FILING DATE: 1998-06-04
; EARLIER APPLICATION NUMBER: 60/048,885
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/049,375
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,881
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,880
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,896
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/049,020
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,876
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,895
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,884
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,894
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,971
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,964
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,882
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,899
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,893
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,900
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,901
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,892
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,915
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/049,019
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,970
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,972
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,916
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/049,373
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,875
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/049,374
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,917
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,949
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,974
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,883
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,897
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,898
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,962
; EARLIER FILING DATE: 1997-06-06
```

EARLIER APPLICATION NUMBER: 60/048,963  
EARLIER FILING DATE: 1997-06-06  
EARLIER APPLICATION NUMBER: 60/048,877  
EARLIER FILING DATE: 1997-06-06  
EARLIER APPLICATION NUMBER: 60/048,878  
EARLIER FILING DATE: 1997-06-06  
EARLIER APPLICATION NUMBER: 60/070,923  
EARLIER FILING DATE: 1997-12-18  
EARLIER APPLICATION NUMBER: 60/092,921  
EARLIER FILING DATE: 1998-07-15  
EARLIER APPLICATION NUMBER: 60/094,657  
EARLIER FILING DATE: 1998-07-30  
NUMBER OF SEQ ID NOS: 1227  
SOFTWARE: PatentIn Ver. 2.0  
SEQ ID NO 381  
LENGTH: 66  
TYPE: PRT  
ORGANISM: Homo sapiens  
FEATURE:  
NAME/KEY: SITE  
LOCATION: (14)  
OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids  
FEATURE:  
NAME/KEY: SITE  
LOCATION: (62)  
OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids  
US-09-205-258-381

Query Match 32.7%; Score 33; DB 4; Length 66;  
Best Local Similarity 35.3%; Pred. No. 2.8e+02;  
Matches 6; Conservative 6; Mismatches 5; Indels 0; Gaps 0;

QY 1 NVLMANNIISKKEIK 17  
| : : : | : | : |  
Db 38 NLSLLTLTKKKKKK 54

Search completed: February 17, 2004, 10:59:40  
Job time : 14.4554 secs

**THIS PAGE BLANK (USPTO)**

Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 NVLMAMNIISK 11  
| | | | | | | | | |  
Db 9 NVLMAMNIISK 19

RESULT 6  
US-09-900-147-16  
; Sequence 16, Application US/09900147  
; Patent No. US20020103121A1  
; GENERAL INFORMATION:  
; APPLICANT: La Thangue, Nicholas B  
; TITLE OF INVENTION: Peptide antagonists of DP transcription factors  
; FILE REFERENCE: 620-67  
; CURRENT APPLICATION NUMBER: US/09/900,147  
; CURRENT FILING DATE: 2001-07-09  
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 09/308,935  
; PRIOR FILING DATE: EARLIER FILING DATE: 1999-05-27  
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: GB 9626589.7  
; PRIOR FILING DATE: EARLIER FILING DATE: 1996-12-20  
; NUMBER OF SEQ ID NOS: 18  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 16  
; LENGTH: 19  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Description of Artificial Sequence: Mutant peptide  
US-09-900-147-16

Query Match 50.5%; Score 51; DB 10; Length 19;  
Best Local Similarity 100.0%; Pred. No. 0.49;  
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 NVLMAMNIISK 11  
| | | | | | | | | |  
Db 9 NVLMAMNIISK 19

RESULT 7  
US-09-900-147-10  
; Sequence 10, Application US/09900147  
; Patent No. US20020103121A1  
; GENERAL INFORMATION:  
; APPLICANT: La Thangue, Nicholas B  
; TITLE OF INVENTION: Peptide antagonists of DP transcription factors  
; FILE REFERENCE: 620-67  
; CURRENT APPLICATION NUMBER: US/09/900,147  
; CURRENT FILING DATE: 2001-07-09  
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 09/308,935  
; PRIOR FILING DATE: EARLIER FILING DATE: 1999-05-27  
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: GB 9626589.7  
; PRIOR FILING DATE: EARLIER FILING DATE: 1996-12-20  
; NUMBER OF SEQ ID NOS: 18  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 10  
; LENGTH: 15  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Description of Artificial Sequence: Synthetic  
US-09-900-147-10

Query Match 49.5%; Score 50; DB 10; Length 15;  
Best Local Similarity 100.0%; Pred. No. 0.54;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 12 EKKEIKWIG 20  
| | | | | | | | | |  
Db 1 EKKEIKWIG 9

RESULT 8  
US-09-764-877-1380  
; Sequence 1380, Application US/09764877  
; Patent No. US20020147140A1  
; GENERAL INFORMATION:  
; APPLICANT: Rosen et al.  
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies  
; FILE REFERENCE: PC005  
; CURRENT APPLICATION NUMBER: US/09/764,877  
; CURRENT FILING DATE: 2001-01-17  
; Prior application data removed - refer to PALM or file wrapper  
; NUMBER OF SEQ ID NOS: 4031  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 1380  
; LENGTH: 29  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-09-764-877-1380

Query Match 49.5%; Score 50; DB 10; Length 29;  
Best Local Similarity 66.7%; Pred. No. 1.1;  
Matches 10; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY 6 MNIISKEKEIKWIG 20  
| | | | | | | | | |  
Db 13 MKIFSKEKKIGWPG 27

RESULT 9  
US-10-242-515-1380  
; Sequence 1380, Application US/10242515  
; Publication No. US20040009488A1  
; GENERAL INFORMATION:  
; APPLICANT: Rosen et al.  
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies  
; FILE REFERENCE: PC005C1  
; CURRENT APPLICATION NUMBER: US/10/242,515  
; CURRENT FILING DATE: 2002-09-13  
; PRIOR APPLICATION NUMBER: 09/764,877  
; PRIOR FILING DATE: 2001-01-17  
; PRIOR APPLICATION NUMBER: 60/179,065  
; PRIOR FILING DATE: 2000-01-31  
; PRIOR APPLICATION NUMBER: 60/180,628  
; PRIOR FILING DATE: 2000-02-04  
; PRIOR APPLICATION NUMBER: 60/214,886  
; PRIOR FILING DATE: 2000-06-28  
; PRIOR APPLICATION NUMBER: 60/217,487  
; PRIOR FILING DATE: 2000-07-11  
; PRIOR APPLICATION NUMBER: 60/225,758  
; PRIOR FILING DATE: 2000-08-14  
; PRIOR APPLICATION NUMBER: 60/220,963  
; PRIOR FILING DATE: 2000-07-26  
; PRIOR APPLICATION NUMBER: 60/217,496  
; PRIOR FILING DATE: 2000-07-11  
; PRIOR APPLICATION NUMBER: 60/225,447  
; PRIOR FILING DATE: 2000-08-14  
; PRIOR APPLICATION NUMBER: 60/218,290  
; PRIOR FILING DATE: 2000-07-14  
; Remaining Prior Application data removed - See File Wrapper or PALM.  
; NUMBER OF SEQ ID NOS: 4031  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 1380  
; LENGTH: 29  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-10-242-515-1380

Query Match 49.5%; Score 50; DB 12; Length 29;  
Best Local Similarity 66.7%; Pred. No. 1.1;  
Matches 10; Conservative 1; Mismatches 4; Indels 0; Gaps 0;





; LENGTH: 19  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Description of Artificial Sequence: Mutant peptide  
US-09-900-147-17

Query Match 44.6%; Score 45; DB 10; Length 19;  
Best Local Similarity 90.9%; Pred. No. 4;  
Matches 10; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 NVLMAMNIISK 11  
Db 9 NVRMAMNIISK 19

## RESULT 14

US-10-214-188-5

; Sequence 5, Application US/10214188

; Publication No. US2003002260A1

; GENERAL INFORMATION:

; APPLICANT: LA THANGUE, NICHOLAS B.

; BERNARDS, RENE

; HIMANS, ELEANORE M.

; TITLE OF INVENTION: TRANSCRIPTION FACTOR E2F-5

; NUMBER OF SEQUENCES: 25

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: NIXON &amp; VANDERHVE P.C.

; STREET: 1100 NORTH GLEBE ROAD

; CITY: ARLINGTON

; STATE: VIRGINIA

; COUNTRY: U.S.A.

; ZIP: 22201-4714

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: Patent In Release #1.0, Version #1.30

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/10/214,188

; FILING DATE: 08-Aug-2002

; CLASSIFICATION: &lt;Unknown&gt;

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: US/08/894,139

; FILING DATE: 13-AUG-1997

; ATTORNEY/AGENT INFORMATION:

; NAME: WILSON, MARY J.

; REGISTRATION NUMBER: 32,955

; REFERENCE/DOCKET NUMBER: 620-22

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: (703) 816-4000

; TELEFAX: (703) 816-4100

; INFORMATION FOR SEQ ID NO: 5:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 74 amino acids

; TYPE: amino acid

; STRANDEDNESS: &lt;Unknown&gt;

; TOPOLOGY: linear

; MOLECULE TYPE: peptide

; SEQUENCE DESCRIPTION: SEQ ID NO: 5:

US-10-214-188-5

Query Match 44.1%; Score 44.5; DB 15; Length 74;  
Best Local Similarity 42.9%; Pred. No. 19;  
Matches 9; Conservative 6; Mismatches 5; Indels 1; Gaps 1;

QY 1 NVLMAMNIISKE-KKEIKWIG 20  
Db 53 NVLEGIQLIAKSKNHIQWLIG 73

## RESULT 15

US-09-932-581-5

; Sequence 5, Application US/09932581  
; Publication No. US20030050264A1  
; GENERAL INFORMATION:  
; APPLICANT: Andrews, William H.  
; APPLICANT: Foster, Christopher A.  
; APPLICANT: Fraser, Stephanie  
; APPLICANT: Mohammadpour, Hamid  
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR MODULATING  
; TITLE OF INVENTION: TELOMERASE REVERSE TRANSCRIPTASE (TERT) EXPRESSION  
; FILE REFERENCE: SIER-005  
; CURRENT APPLICATION NUMBER: US/09/932,581  
; CURRENT FILING DATE: 2001-08-17  
; PRIOR APPLICATION NUMBER: 60/227,865  
; PRIOR FILING DATE: 2000-08-24  
; PRIOR APPLICATION NUMBER: 60/230,174  
; PRIOR FILING DATE: 2000-09-01  
; PRIOR APPLICATION NUMBER: 60/238,345  
; PRIOR FILING DATE: 2000-10-05  
; NUMBER OF SEQ ID NOS: 25  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 5  
; LENGTH: 85  
; TYPE: PRT  
; ORGANISM: human  
US-09-932-581-5

Query Match 44.1%; Score 44.5; DB 11; Length 85;  
Best Local Similarity 42.9%; Pred. No. 22;  
Matches 9; Conservative 6; Mismatches 5; Indels 1; Gaps 1;

QY 1 NVLMAMNIISKE-KKEIKWIG 20

Db 63 NVLEGIQLIAKSKNHIQWLIG 83

## RESULT 16

US-10-338-294-5

; Sequence 5, Application US/10338294

; Publication No. US20030171326A1

; GENERAL INFORMATION:

; APPLICANT: Andrews, William H.

; APPLICANT: Foster, Christopher A.

; APPLICANT: Fraser, Stephanie

; APPLICANT: Mohammadpour, Hamid

; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR MODULATING

; TITLE OF INVENTION: TELOMERASE REVERSE TRANSCRIPTASE (TERT) EXPRESSION

; FILE REFERENCE: SIER-005

; CURRENT APPLICATION NUMBER: US/10/338,294

; CURRENT FILING DATE: 2003-01-07

; PRIOR APPLICATION NUMBER: US/09/932,581

; PRIOR FILING DATE: 2001-08-17

; PRIOR APPLICATION NUMBER: 60/227,865

; PRIOR FILING DATE: 2000-08-24

; PRIOR APPLICATION NUMBER: 60/230,174

; PRIOR FILING DATE: 2000-09-01

; PRIOR APPLICATION NUMBER: 60/238,345

; PRIOR FILING DATE: 2000-10-05

; NUMBER OF SEQ ID NOS: 25

; SOFTWARE: FastSeq for Windows Version 4.0

; SEQ ID NO 5

; LENGTH: 85

; TYPE: PRT

; ORGANISM: human

US-10-338-294-5

Query Match 44.1%; Score 44.5; DB 12; Length 85;  
Best Local Similarity 42.9%; Pred. No. 22;  
Matches 9; Conservative 6; Mismatches 5; Indels 1; Gaps 1;

QY 1 NVLMAMNIISKE-KKEIKWIG 20

Db 63 NVLEGIQLIAKSKNHIQWLIG 83

```

; MOLECULE TYPE: peptide
; SEQUENCE DESCRIPTION: SEQ ID NO: 7:
US-10-214-188-7

Query Match      42.1%; Score 42.5; DB 15; Length 74;
Best Local Similarity 38.1%; Pred. No. 38;
Matches 8; Conservative 7; Mismatches 5; Indels 1; Gaps 1;

QY 1 NVLMAMNIIISKE-KKEIKWIG 20
   ||| :||| :| :|||
Db 53 NVLEGIHLIKKSKNHVQMVG 73

RESULT 19
US-09-900-147-2
; Sequence 2, Application US/09900147
; Patent No. US20020103121A1
; GENERAL INFORMATION:
; APPLICANT: La Thangue, Nicholas B
; APPLICANT: Bandara, Lagantha R
; TITLE OF INVENTION: Peptide antagonists of DP transcription factors
; FILE REFERENCE: 620-67
; CURRENT APPLICATION NUMBER: US/09/900,147
; CURRENT FILING DATE: 2001-07-09
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 09/308,935
; PRIOR FILING DATE: EARLIER FILING DATE: 1999-05-27
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: GB 9626589.7
; PRIOR FILING DATE: EARLIER FILING DATE: 1996-12-20
; NUMBER OF SEQ ID NOS: 18
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2
; LENGTH: 9
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Synthetic
US-09-900-147-2

Query Match      41.6%; Score 42; DB 10; Length 9;
Best Local Similarity 100.0%; Pred. No. 7e+05;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 NVLMAMNIII 9
   ||| |||||
Db 1 NVLMAMNIII 9

RESULT 20
US-10-097-065-532
; Sequence 532, Application US/10097065
; Publication No. US20030055236A1
; GENERAL INFORMATION:
; APPLICANT: Moore, Paul A. et al.
; TITLE OF INVENTION: 110 Human Secreted Proteins
; FILE REFERENCE: P2021P1
; CURRENT APPLICATION NUMBER: US/10/097,065
; CURRENT FILING DATE: 2002-03-14
; PRIOR APPLICATION NUMBER: PCT/US98/27059
; PRIOR FILING DATE: 1998-12-17
; PRIOR APPLICATION NUMBER: 60/070,923
; PRIOR FILING DATE: 1997-12-18
; PRIOR APPLICATION NUMBER: 60/068,007
; PRIOR FILING DATE: 1997-12-18
; PRIOR APPLICATION NUMBER: 60/068,057
; PRIOR FILING DATE: 1997-12-18
; PRIOR APPLICATION NUMBER: 60/068,006
; PRIOR FILING DATE: 1997-12-18
; PRIOR APPLICATION NUMBER: 60/068,369
; PRIOR FILING DATE: 1997-12-19
; PRIOR APPLICATION NUMBER: 60/068,367
; PRIOR FILING DATE: 1997-12-19
; PRIOR APPLICATION NUMBER: 60/068,368
; PRIOR FILING DATE: 1997-12-19

```

```
; PRIOR APPLICATION NUMBER: 60/068,169
; PRIOR FILING DATE: 1997-12-19
; PRIOR APPLICATION NUMBER: 60/068,053
; PRIOR FILING DATE: 1997-12-18
; PRIOR APPLICATION NUMBER: 60/068,064
; PRIOR FILING DATE: 1997-12-18
; PRIOR APPLICATION NUMBER: 60/068,054
; PRIOR FILING DATE: 1997-12-18
; PRIOR APPLICATION NUMBER: 60/068,008
; PRIOR FILING DATE: 1997-12-18
; PRIOR APPLICATION NUMBER: 60/068,365
; PRIOR FILING DATE: 1997-12-19
; NUMBER OF SEQ ID NOS: 672
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 532
; LENGTH: 57
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-097-065-532

Query Match          39.6%; Score 40; DB 15; Length 57;
Best Local Similarity 31.6%; Pred. No. 70;
Matches 6; Conservative 6; Mismatches 7; Indels 0; Gaps 0;

QY      1 NVLMANNIISKEKEIKWI 19
Db      1 NFCVSKTNFVRKPKW 19

RESULT 21
US-09-864-408A-3370
; Sequence 3370, Application US/09864408A
; Publication No. US20040009474A1
; GENERAL INFORMATION:
; APPLICANT: Leach, Martin D.
; APPLICANT: Shimkets, Richard A.
; TITLE OF INVENTION: No. US20040009474A1 Human Polynucleotides and Polypeptides Encod
; FILE REFERENCE: 21402-012
; CURRENT APPLICATION NUMBER: US/09/864,408A
; CURRENT FILING DATE: 2001-05-24
; PRIOR APPLICATION NUMBER: 60/206,690
; PRIOR FILING DATE: 2000-05-24
; NUMBER OF SEQ ID NOS: 9068
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 3370
; LENGTH: 63
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)..(1)
; OTHER INFORMATION: Wherein Xaa may be any naturally occurring amino acid
US-09-864-408A-3370

Query Match          39.6%; Score 40; DB 12; Length 63;
Best Local Similarity 50.0%; Pred. No. 77;
Matches 6; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

QY      9 ISKEKEIKWIG 20
Db      48 VQKESKENOWLG 59

RESULT 22
US-09-864-761-35555
; Sequence 35555, Application US/09864761
; Patent No. US20020048763A1
; GENERAL INFORMATION:
; APPLICANT: Penn, Sharon G.
; APPLICANT: Rank, David R.
; APPLICANT: Hanzel, David K.
; APPLICANT: Chen, Wensheng
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
```

```
; TITLE OF INVENTION: GENE EXPRESSION ANALYSIS BY MICROARRAY
; FILE REFERENCE: Aeomica-X-1
; CURRENT APPLICATION NUMBER: US/09/864,761
; CURRENT FILING DATE: 2001-05-23
; PRIOR APPLICATION NUMBER: US 60/180,312
; PRIOR FILING DATE: 2000-02-04
; PRIOR APPLICATION NUMBER: US 60/207,456
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: US 09/632,366
; PRIOR FILING DATE: 2000-08-03
; PRIOR APPLICATION NUMBER: GB 24263.6
; PRIOR FILING DATE: 2000-10-04
; PRIOR APPLICATION NUMBER: US 60/236,359
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: PCT/US01/00666
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00667
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00664
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00669
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00665
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00668
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00663
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00662
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00661
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00670
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: US 60/234,687
; PRIOR FILING DATE: 2000-09-21
; PRIOR APPLICATION NUMBER: US 09/608,408
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: US 09/774,203
; PRIOR FILING DATE: 2001-01-29
; NUMBER OF SEQ ID NOS: 49117
; SOFTWARE: Annomax Sequence Listing Engine vers. 1.1
; SEQ ID NO 35555
; LENGTH: 75
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: MAP TO AL109953.6
; OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 1.9
; OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 1.5
; OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 2.5
; OTHER INFORMATION: EXPRESSED IN BT474, SIGNAL = 3.5
; OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 21
; OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 2.2
; OTHER INFORMATION: EXPRESSED IN HBL100, SIGNAL = 3.7
; OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 3.4
; OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 2
; OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 2.4
; OTHER INFORMATION: EST HUMAN HIT: A1817324.1, EVALUE 4.00e-22
; OTHER INFORMATION: SWISSPROT HIT: P51578, EVALUE 2.50e+00
US-09-864-761-35555

Query Match          39.6%; Score 40; DB 9; Length 75;
Best Local Similarity 50.0%; Pred. No. 93;
Matches 6; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

QY      9 ISKEKEIKWIG 20
Db      31 ITKEKSSLRWAG 42

RESULT 23
US-10-214-188-9
```

```

Query Match      38.6%; Score 19; DB 12; Length 84;
Best Local Similarity 42.1%; Pred. No. 1.5e+02;
Matches      8; Conservative      5; Mismatches      4; Indels      2; Gaps
QY      3 LMANNIISKEKEIK-WI 19
      |:|::|::|:|
Db      56 LIARRVVGREAKEIERYWI 74

RESULT 25
US-09-932-581-6
; Sequence 6, Application US/09932581
; Publication No. US20030050264A1
; GENERAL INFORMATION:
; APPLICANT: Andrews, William H.
; APPLICANT: Foster, Christopher A.
; APPLICANT: Fraser, Stephanie
; APPLICANT: Mohammadpour, Hamid
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR MODULATING
; TITLE OF INVENTION: TELOMERASE REVERSE TRANSCRIPTASE (TERT)
; FILE REFERENCE: SIER-005
; CURRENT APPLICATION NUMBER: US/09/932,581
; CURRENT FILING DATE: 2001-08-17
; PRIOR APPLICATION NUMBER: 60/227,865
; PRIOR FILING DATE: 2000-08-24
; PRIOR APPLICATION NUMBER: 60/230,174
; PRIOR FILING DATE: 2000-09-01
; PRIOR APPLICATION NUMBER: 60/238,345
; PRIOR FILING DATE: 2000-10-05
; NUMBER OF SEQ ID NOS: 25
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 6
; LENGTH: 76
; TYPE: PRT
; ORGANISM: human
US-09-932-581-6

Query Match      38.1%; Score 38.5; DB 11; Length 76;
Best Local Similarity 42.9%; Pred. No. 1.6e+02;
Matches      9; Conservative      4; Mismatches      7; Indels      1; Gaps
QY      1 NVLMANNIISKE-KKEIKWIG 20
      |||::|::|:|:|
Db      53 NVLEGIGLIEKKSQNSIQWK 73

RESULT 26
US-10-338-294-6
; Sequence 6, Application US/10338294
; Publication No. US20030171326A1
; GENERAL INFORMATION:
; APPLICANT: Andrews, William H.
; APPLICANT: Foster, Christopher A.
; APPLICANT: Fraser, Stephanie
; APPLICANT: Mohammadpour, Hamid
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR MODULATING
; TITLE OF INVENTION: TELOMERASE REVERSE TRANSCRIPTASE (TERT)
; FILE REFERENCE: SIER-005
; CURRENT APPLICATION NUMBER: US/10/338,294
; CURRENT FILING DATE: 2003-01-07
; PRIOR APPLICATION NUMBER: US/09/932,581
; PRIOR FILING DATE: 2001-08-17
; PRIOR APPLICATION NUMBER: 60/227,865
; PRIOR FILING DATE: 2000-08-24
; PRIOR APPLICATION NUMBER: 60/230,174
; PRIOR FILING DATE: 2000-09-01
; PRIOR APPLICATION NUMBER: 60/238,345
; PRIOR FILING DATE: 2000-10-05
; NUMBER OF SEQ ID NOS: 25
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 6
; LENGTH: 76
; TYPE: PRT

```

; ORGANISM: human  
US-10-338-294-6

Query Match 38.1%; Score 38.5; DB 12; Length 76;  
Best Local Similarity 42.9%; Pred. No. 1.6e+02;  
Matches 9; Conservative 4; Mismatches 7; Indels 1; Gaps 1;

QY 1 NVLMAMNIISKE-KKEIKWIG 20  
||| : : : : :  
Db 53 NVLEGIGLIEKSKNSIQWKG 73

## RESULT 27

US-10-029-386-29071  
; Sequence 29071, Application US/10029386  
; Publication No. US20030194704A1  
; GENERAL INFORMATION:  
; APPLICANT: Penn, Sharron G.  
; APPLICANT: Rank, David R.  
; APPLICANT: Hanzel, David K.  
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR  
; FILE REFERENCE: AEOMICA-X-2  
; CURRENT APPLICATION NUMBER: US/10/029,386  
; CURRENT FILING DATE: 2001-12-20  
; NUMBER OF SEQ ID NOS: 34288  
; SOFTWARE: Annomax Sequence Listing Engine vers. 1.1  
; SEQ ID NO 29071  
; LENGTH: 76  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
; FEATURE:  
; OTHER INFORMATION: MAP TO CHR20.1  
; OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 6.1  
; OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 5.4  
; OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 4.6  
; OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 3.3  
; OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 4.8  
; OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 4.4  
; OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 4  
; OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 5.4  
; OTHER INFORMATION: SWISSPROT HIT: Q01094, EVALUE 1.00e-36  
US-10-029-386-29071

Query Match 38.1%; Score 38.5; DB 12; Length 76;  
Best Local Similarity 40.0%; Pred. No. 1.6e+02;  
Matches 8; Conservative 6; Mismatches 5; Indels 1; Gaps 1;

QY 1 NVLMAMNIISKE-KKEIKWI 19  
||| : : : : :  
Db 57 NVLEGIGLIAKSKNHIQWL 76

## RESULT 28

US-10-165-614-3  
; Sequence 3, Application US/10165614  
; Publication No. US20020193289A1  
; GENERAL INFORMATION:  
; APPLICANT: Andrews, William H.  
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR MODULATING  
; FILE REFERENCE: SIER-018  
; CURRENT APPLICATION NUMBER: US/10/165,614  
; CURRENT FILING DATE: 2002-06-25  
; PRIOR APPLICATION NUMBER: 60/296,992  
; PRIOR FILING DATE: 2001-06-07  
; NUMBER OF SEQ ID NOS: 4  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 3  
; LENGTH: 76  
; TYPE: PRT  
; ORGANISM: homo sapiens  
US-10-165-614-3

Query Match 38.1%; Score 38.5; DB 14; Length 76;  
Best Local Similarity 42.9%; Pred. No. 1.6e+02;  
Matches 9; Conservative 4; Mismatches 7; Indels 1; Gaps 1;

QY 1 NVLMAMNIISKE-KKEIKWIG 20  
||| : : : : :  
Db 53 NVLEGIGLIEKSKNSIQWKG 73

## RESULT 29

US-09-864-761-43984  
; Sequence 43984, Application US/09864761  
; Patent No. US20020048763A1  
; GENERAL INFORMATION:  
; APPLICANT: Penn, Sharron G.  
; APPLICANT: Rank, David R.  
; APPLICANT: Hanzel, David K.  
; APPLICANT: Chen, Wensheng  
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR  
; FILE REFERENCE: Aeomica-X-1  
; CURRENT APPLICATION NUMBER: US/09/864,761  
; CURRENT FILING DATE: 2001-05-23  
; PRIOR APPLICATION NUMBER: US 60/180,312  
; PRIOR FILING DATE: 2000-02-04  
; PRIOR APPLICATION NUMBER: US 60/207,456  
; PRIOR FILING DATE: 2000-05-26  
; PRIOR APPLICATION NUMBER: US 09/632,366  
; PRIOR FILING DATE: 2000-08-03  
; PRIOR APPLICATION NUMBER: GB 24263.6  
; PRIOR FILING DATE: 2000-10-04  
; PRIOR APPLICATION NUMBER: US 60/236,359  
; PRIOR FILING DATE: 2000-09-27  
; PRIOR APPLICATION NUMBER: PCT/US01/00666  
; PRIOR FILING DATE: 2001-01-30  
; PRIOR APPLICATION NUMBER: PCT/US01/00667  
; PRIOR FILING DATE: 2001-01-30  
; PRIOR APPLICATION NUMBER: PCT/US01/00664  
; PRIOR FILING DATE: 2001-01-30  
; PRIOR APPLICATION NUMBER: PCT/US01/00669  
; PRIOR FILING DATE: 2001-01-30  
; PRIOR APPLICATION NUMBER: PCT/US01/00665  
; PRIOR FILING DATE: 2001-01-30  
; PRIOR APPLICATION NUMBER: PCT/US01/00668  
; PRIOR FILING DATE: 2001-01-30  
; PRIOR APPLICATION NUMBER: PCT/US01/00663  
; PRIOR FILING DATE: 2001-01-30  
; PRIOR APPLICATION NUMBER: PCT/US01/00662  
; PRIOR FILING DATE: 2001-01-30  
; PRIOR APPLICATION NUMBER: PCT/US01/00661  
; PRIOR FILING DATE: 2001-01-30  
; PRIOR APPLICATION NUMBER: PCT/US01/00670  
; PRIOR FILING DATE: 2001-01-30  
; PRIOR APPLICATION NUMBER: US 60/234,687  
; PRIOR FILING DATE: 2000-09-21  
; PRIOR APPLICATION NUMBER: US 09/608,408  
; PRIOR FILING DATE: 2000-06-30  
; PRIOR APPLICATION NUMBER: US 09/774,203  
; PRIOR FILING DATE: 2001-01-29  
; NUMBER OF SEQ ID NOS: 49117  
; SOFTWARE: Annomax Sequence Listing Engine vers. 1.1  
; SEQ ID NO 43984  
; LENGTH: 31  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
; FEATURE:  
; OTHER INFORMATION: MAP TO AC005386.1  
; OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 0.68  
; OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 0.79  
; OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 0.86  
; OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 0.76  
; OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 0.74

```
; OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 0.72
US-09-864-761-43984

Query Match      37.6%; Score 38; DB 9; Length 31;
Best Local Similarity 50.0%; Pred. No. 75;
Matches 6; Conservative 4; Mismatches 2; Indels 0; Gaps 0;

QY 7 NISKEKEIKW 18
Db 7 NMISKOSRKKKW 18

RESULT 30
US-10-083-357-694
; Sequence 694, Application US/10083357
; Publication No. US20030054370A1
; GENERAL INFORMATION:
; APPLICANT: Qiandong Zeng et al.
; TITLE OF INVENTION: Systemic Discovery of New Genes
; FILE REFERENCE: 032796-090
; CURRENT APPLICATION NUMBER: US/10/083,357
; CURRENT FILING DATE: 2002-02-27
; NUMBER OF SEQ ID NOS: 1346
; SEQ ID NO 694
; LENGTH: 69
; TYPE: PRT
; ORGANISM: Saccharomyces cerevisiae
US-10-083-357-694

Query Match      37.6%; Score 38; DB 15; Length 69;
Best Local Similarity 40.9%; Pred. No. 1.7e+02;
Matches 9; Conservative 3; Mismatches 6; Indels 4; Gaps 1;

QY 3 LMANNIISKE-----KKEIKWIG 20
Db 34 LVCNRIIKQSITTKGKQWFG 55

RESULT 31
US-10-029-386-30735
; Sequence 30735, Application US/10029386
; Publication No. US20030194704A1
; GENERAL INFORMATION:
; APPLICANT: Penn, Sharon G.
; APPLICANT: Rank, David R.
; APPLICANT: Hanzel, David K.
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
; FILE REFERENCE: ABOICA-X-2
; CURRENT APPLICATION NUMBER: US/10/029,386
; CURRENT FILING DATE: 2001-12-20
; NUMBER OF SEQ ID NOS: 34288
; SOFTWARE: Annomax Sequence Listing Engine vers. 1.1
; SEQ ID NO 30735
; LENGTH: 81
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: MAP TO CHR13.1
; OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 0.7
; OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 0.61
; OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 0.75
; OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 0.62
; OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 0.59
; OTHER INFORMATION: SWISSPROT HIT: P07271, EVALUATE 2.20e-01
US-10-029-386-30735

Query Match      37.6%; Score 38; DB 12; Length 81;
Best Local Similarity 46.2%; Pred. No. 2e+02;
Matches 6; Conservative 4; Mismatches 3; Indels 0; Gaps 0;

QY 7 NISKEKEIKWI 19
Db 7 NMISKOSRKKKW 18

; OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 0.72
US-09-864-761-43984

Query Match      37.6%; Score 38; DB 9; Length 31;
Best Local Similarity 50.0%; Pred. No. 75;
Matches 6; Conservative 4; Mismatches 2; Indels 0; Gaps 0;

QY 7 NISKEKEIKW 18
Db 7 NMISKOSRKKKW 18

RESULT 32
US-09-916-790-25
; Sequence 25, Application US/09916790
; Patent No. US20020061573A1
; GENERAL INFORMATION:
; APPLICANT: Meyers, Rachael
; APPLICANT: Kapeller-Libermann, Rosana
; APPLICANT: Silos-Santiago, Immaculada
; TITLE OF INVENTION: 18431 AND 32374, NOVEL HUMAN PROTEIN
; FILE REFERENCE: 381552002700
; CURRENT APPLICATION NUMBER: US/09/916,790
; CURRENT FILING DATE: 2001-07-27
; PRIOR APPLICATION NUMBER: 60/221,543
; PRIOR FILING DATE: 2000-07-28
; NUMBER OF SEQ ID NOS: 37
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 25
; LENGTH: 85
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Consensus amino acid sequence
US-09-916-790-25

Query Match      37.6%; Score 38; DB 9; Length 85;
Best Local Similarity 41.2%; Pred. No. 2.1e+02;
Matches 7; Conservative 5; Mismatches 5; Indels 0; Gaps 0;

QY 1 NVLMANNIISKEKEIK 17
Db 57 NILLDNMMVAKGDSSEIK 73

RESULT 33
US-09-922-138-20
; Sequence 20, Application US/09922138
; Patent No. US20020061574A1
; GENERAL INFORMATION:
; APPLICANT: Millennium Pharmaceuticals, Inc.
; APPLICANT: Meyers, Rachel
; APPLICANT: Silos-Santiago, Immaculada
; TITLE OF INVENTION: 16658, 14223, AND 16002, NOVEL HUMAN
; FILE REFERENCE: 38155-20030.00
; CURRENT APPLICATION NUMBER: US/09/922,138
; CURRENT FILING DATE: 2001-11-28
; PRIOR APPLICATION NUMBER: US 60/229,299
; PRIOR FILING DATE: 2000-09-01
; NUMBER OF SEQ ID NOS: 38
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 20
; LENGTH: 100
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Consensus amino acid sequence
US-09-922-138-20

Query Match      37.6%; Score 38; DB 9; Length 100;
Best Local Similarity 41.2%; Pred. No. 2.5e+02;
Matches 7; Conservative 5; Mismatches 5; Indels 0; Gaps 0;

QY 1 NVLMANNIISKEKEIK 17
Db 57 NILLDNMMVAKGDSSEIK 73

RESULT 34
US-09-867-550-1920
```

```
; Sequence 1920, Application US/09867550
; Patent No. US20020082206A1
; GENERAL INFORMATION:
; APPLICANT: Leach, Martin D.
; APPLICANT: Mehraban, Fuad,
; APPLICANT: Conley, Pamela
; APPLICANT: Law, Debbie
; APPLICANT: Topper, James
; TITLE OF INVENTION: No. US20020082206A1 Polynucleotides from Atherogenic Cells and
; FILE REFERENCE: 21402-013 (Cura-313)
; CURRENT APPLICATION NUMBER: US/09/867,550
; CURRENT FILING DATE: 2001-09-20
; PRIOR APPLICATION NUMBER: USSN 60/208,427
; PRIOR FILING DATE: 2000-05-30
; NUMBER OF SEQ ID NOS: 2125
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1920
; LENGTH: 77
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-867-550-1920

Query Match          35.1%; Score 35.5; DB 9; Length 77;
Best Local Similarity 38.9%; Pred. No. 4.6e+02;
Matches 7; Conservative 7; Mismatches 3; Indels 1; Gaps 1;

QY      4 MAMNIIS-KEKKEIKWIG 20
DB      44 ISKSVISLLEKRLPWIG 61

RESULT 35
US-09-864-761-33458
; Sequence 33458, Application US/09864761
; Patent No. US20020048763A1
; GENERAL INFORMATION:
; APPLICANT: Penn, Sharron G.
; APPLICANT: Rank, David R.
; APPLICANT: Hanzel, David K.
; APPLICANT: Chen, Wensheng
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
; FILE REFERENCE: Aeomica-X-1
; CURRENT APPLICATION NUMBER: US/09/864,761
; CURRENT FILING DATE: 2001-05-23
; PRIOR APPLICATION NUMBER: US 60/180,312
; PRIOR FILING DATE: 2000-02-04
; PRIOR APPLICATION NUMBER: US 60/207,456
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: US 09/632,366
; PRIOR FILING DATE: 2000-08-03
; PRIOR APPLICATION NUMBER: GB 24263.6
; PRIOR FILING DATE: 2000-10-04
; PRIOR APPLICATION NUMBER: US 60/236,359
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: PCT/US01/00666
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00667
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00664
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00669
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00665
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00668
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00663
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00662
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00661
```

```
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00670
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: US 60/234,687
; PRIOR FILING DATE: 2000-09-21
; PRIOR APPLICATION NUMBER: US 09/608,408
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: US 09/774,203
; PRIOR FILING DATE: 2001-01-29
; NUMBER OF SEQ ID NOS: 49117
; SOFTWARE: Anomax Sequence Listing Engine vers. 1.1
; SEQ ID NO 33458
; LENGTH: 56
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: MAP TO AC007842.1
; OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 0.85
; OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 1
; OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 1.1
; OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 1.5
; OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 0.7
; OTHER INFORMATION: EXPRESSED IN BT474, SIGNAL = 1.4
; OTHER INFORMATION: EXPRESSED IN HBL100, SIGNAL = 1.2
; OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 0.85
; OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 1
; OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 1.3
; OTHER INFORMATION: EST HUMAN HIT: AA084272.1, EVALUE 7.00e-04
; OTHER INFORMATION: SWISSPROT HIT: P15870, EVALUE 4.50e+00
US-09-864-761-33458

Query Match          34.7%; Score 35; DB 9; Length 58;
Best Local Similarity 23.5%; Pred. No. 4.1e+02;
Matches 4; Conservative 10; Mismatches 3; Indels 0; Gaps 0;

QY      2 VLMAMNIISKEKEIKW 18
DB      36 MVCNVNMLRKDRIRW 52

RESULT 36
US-09-864-408A-4500
; Sequence 4500, Application US/09864408A
; Publication No. US20040009474A1
; GENERAL INFORMATION:
; APPLICANT: Leach, Martin D.
; APPLICANT: Shimkets, Richard A.
; TITLE OF INVENTION: No. US20040009474A1 Human Polynucleotides and Polypeptides Encod
; FILE REFERENCE: 21402-012
; CURRENT APPLICATION NUMBER: US/09/864,408A
; CURRENT FILING DATE: 2001-05-24
; PRIOR APPLICATION NUMBER: 60/206,690
; PRIOR FILING DATE: 2000-05-24
; NUMBER OF SEQ ID NOS: 9068
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 4500
; LENGTH: 98
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-864-408A-4500

Query Match          34.7%; Score 35; DB 12; Length 98;
Best Local Similarity 33.3%; Pred. No. 7e+02;
Matches 6; Conservative 5; Mismatches 7; Indels 0; Gaps 0;

QY      3 LMAMNIISKEKEIKWIG 20
DB      12 LGALTIVLSCALREYRTWG 29

RESULT 37
US-09-900-147-11
; Sequence 11, Application US/09900147
```

```
; Patent No. US20020103121A1
; GENERAL INFORMATION:
; APPLICANT: La Thangue, Nicholas B
; APPLICANT: Bandara, Lasantha R
; TITLE OF INVENTION: Peptide antagonists of DP transcription factors
; FILE REFERENCE: 620-67
; CURRENT APPLICATION NUMBER: US/09/900,147
; PRIOR FILING DATE: 2001-07-09
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 09/308,935
; PRIOR FILING DATE: EARLIER FILING DATE: 1999-05-27
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: GB 9626589.7
; PRIOR FILING DATE: EARLIER FILING DATE: 1996-12-20
; NUMBER OF SEQ ID NOS: 18
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 11
; LENGTH: 14
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Synthetic
; US-09-900-147-11

Query Match          33.7%; Score 34; DB 10; Length 14;
Best Local Similarity 100.0%; Pred. No. 1.3e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 NVLMANN 7
   |||||
Db 8 NVLMANN 14

RESULT 38
US-10-157-031-192
; Sequence 192, Application US/10157031
; Publication No. US20030108890A1
; GENERAL INFORMATION:
; APPLICANT: Baranova, A. V.
; APPLICANT: Yankovsky, N. K.
; APPLICANT: Kozlov, A. P.
; APPLICANT: Lobashev, A. V.
; APPLICANT: Krukovskaya, L. L.
; TITLE OF INVENTION: In silico screening for phenotype-associated expressed sequences
; FILE REFERENCE: 2760-103
; CURRENT APPLICATION NUMBER: US/10/157,031
; CURRENT FILING DATE: 2002-05-30
; NUMBER OF SEQ ID NOS: 415
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 192
; LENGTH: 18
; TYPE: PRT
; ORGANISM: Homo sapiens
; US-10-157-031-192

Query Match          33.7%; Score 34; DB 15; Length 18;
Best Local Similarity 42.9%; Pred. No. 1.7e+02;
Matches 6; Conservative 5; Mismatches 3; Indels 0; Gaps 0;

QY 4 MAMNIISKEKKEIK 17
   |::|::|
Db 1 MYVNVLQKKKKKK 14

RESULT 39
US-09-864-761-48415
; Sequence 48415, Application US/09864761
; Patent No. US20020048763A1
; GENERAL INFORMATION:
; APPLICANT: Penn, Sharon G.
; APPLICANT: Rank, David R.
; APPLICANT: Hanzel, David K.
; APPLICANT: Chen, Wensheng
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
; TITLE OF INVENTION: GENE EXPRESSION ANALYSIS BY MICROARRAY
```

```
; FILE REFERENCE: Aomica-X-1
; CURRENT APPLICATION NUMBER: US/09/864,761
; CURRENT FILING DATE: 2001-05-23
; PRIOR APPLICATION NUMBER: US 60/180,312
; PRIOR FILING DATE: 2000-02-04
; PRIOR APPLICATION NUMBER: US 60/207,456
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: US 09/632,366
; PRIOR FILING DATE: 2000-08-03
; PRIOR APPLICATION NUMBER: GB 24263.6
; PRIOR FILING DATE: 2000-10-04
; PRIOR APPLICATION NUMBER: US 60/236,359
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: PCT/US01/006666
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/006667
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/006664
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/006669
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/006665
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/006668
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/006663
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/006662
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: US 60/234,687
; PRIOR FILING DATE: 2000-09-21
; PRIOR APPLICATION NUMBER: US 09/608,408
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: US 09/774,203
; PRIOR FILING DATE: 2001-01-29
; NUMBER OF SEQ ID NOS: 49117
; SOFTWARE: Anomax Sequence Listing Engine vers. 1.1
; SEQ ID NO 48415
; LENGTH: 32
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: MAP TO AL096704.6
; OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 6.1
; OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 0.92
; OTHER INFORMATION: EST_HUMAN HIT: AA250859.1, EVALUATE 7.80e-02
; US-09-864-761-48415

Query Match          33.7%; Score 34; DB 9; Length 32;
Best Local Similarity 50.0%; Pred. No. 3.1e+02;
Matches 8; Conservative 4; Mismatches 2; Indels 2; Gaps 1;

QY 2 VLMAMNIISKEKKEIK 17
   |::|::|::|
Db 14 ILQQIN--SKEKKQMK 27

RESULT 40
US-09-764-869-723
; Sequence 723, Application US/09764869
; Patent No. US20020061521A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PC007
; CURRENT APPLICATION NUMBER: US/09/764,869
; CURRENT FILING DATE: 2001-01-17
; Prior application data removed - refer to PALM or file wrapper
; NUMBER OF SEQ ID NOS: 2442
```



; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 723  
; LENGTH: 35  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: SITE  
; LOCATION: (22)  
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids  
US-09-764-869-723

Query Match 33.7%; Score 34; DB 9; Length 35;  
Best Local Similarity 35.7%; Pred. No. 3.4e+02;  
Matches 5; Conservative 6; Mismatches 3; Indels 0; Gaps 0;

QY 5 AMNIISEKKEIKW 18  
|:::|:|:|:  
Db 7 AIHLSDKKKEKRW 20

## RESULT 41

US-10-227-577-723  
; Sequence 723, Application US/10227577  
; Publication No. US20040005575A1  
; GENERAL INFORMATION:  
; APPLICANT: Rosen et al.  
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies  
; FILE REFERENCE: PC007C2  
; CURRENT APPLICATION NUMBER: US/10/227,577  
; CURRENT FILING DATE: 2002-08-26

; PRIOR APPLICATION NUMBER: 10/091,504  
; PRIOR FILING DATE: 2002-03-07  
; PRIOR APPLICATION NUMBER: 09/764,869  
; PRIOR FILING DATE: 2001-01-17  
; PRIOR APPLICATION NUMBER: 60/179,065  
; PRIOR FILING DATE: 2000-01-31  
; PRIOR APPLICATION NUMBER: 60/180,628  
; PRIOR FILING DATE: 2000-02-04  
; PRIOR APPLICATION NUMBER: 60/214,886  
; PRIOR FILING DATE: 2000-06-28  
; PRIOR APPLICATION NUMBER: 60/217,487  
; PRIOR FILING DATE: 2000-07-11  
; PRIOR APPLICATION NUMBER: 60/225,758  
; PRIOR FILING DATE: 2000-08-14  
; PRIOR APPLICATION NUMBER: 60/220,963  
; PRIOR FILING DATE: 2000-07-26  
; PRIOR APPLICATION NUMBER: 60/217,496  
; PRIOR FILING DATE: 2000-07-11  
; PRIOR APPLICATION NUMBER: 60/225,447  
; PRIOR FILING DATE: 2000-08-14  
; Remaining prior Application data removed - See File Wrapper or PALM.  
; NUMBER OF SEQ ID NOS: 2442  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 723  
; LENGTH: 35  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: misc\_feature  
; LOCATION: (22)  
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids  
US-10-227-577-723

Query Match 33.7%; Score 34; DB 12; Length 35;  
Best Local Similarity 35.7%; Pred. No. 3.4e+02;  
Matches 5; Conservative 6; Mismatches 3; Indels 0; Gaps 0;

QY 5 AMNIISEKKEIKW 18  
|:::|:|:|:  
Db 7 AIHLSDKKKEKRW 20

## RESULT 42

US-10-091-504-723  
; Sequence 723, Application US/10091504  
; Publication No. US20030059908A1  
; GENERAL INFORMATION:  
; APPLICANT: Rosen et al.  
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies  
; FILE REFERENCE: PC007C1  
; CURRENT APPLICATION NUMBER: US/10/091,504  
; CURRENT FILING DATE: 2002-03-07  
; NUMBER OF SEQ ID NOS: 2442  
; Prior Application removed - See File Wrapper or PALM  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 723  
; LENGTH: 35  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: misc\_feature  
; LOCATION: (22)  
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids  
US-10-091-504-723

Query Match 33.7%; Score 34; DB 15; Length 35;  
Best Local Similarity 35.7%; Pred. No. 3.4e+02;  
Matches 5; Conservative 6; Mismatches 3; Indels 0; Gaps 0;

QY 5 AMNIISEKKEIKW 18  
|:::|:|:|:  
Db 7 AIHLSDKKKEKRW 20

## RESULT 43

US-09-864-761-47714  
; Sequence 47714, Application US/09864761  
; Patent No. US20020048763A1  
; GENERAL INFORMATION:  
; APPLICANT: Penn, Sharron G.  
; APPLICANT: Rank, David R.  
; APPLICANT: Hanzel, David K.  
; APPLICANT: Chen, Wensheng  
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR  
; FILE REFERENCE: Aeomica-X-1  
; CURRENT APPLICATION NUMBER: US/09/864,761  
; CURRENT FILING DATE: 2001-05-23  
; PRIOR APPLICATION NUMBER: US 60/180,312  
; PRIOR FILING DATE: 2000-02-04  
; PRIOR APPLICATION NUMBER: US 60/207,456  
; PRIOR FILING DATE: 2000-05-26  
; PRIOR APPLICATION NUMBER: US 09/632,366  
; PRIOR FILING DATE: 2000-08-03  
; PRIOR APPLICATION NUMBER: GB 24263.6  
; PRIOR FILING DATE: 2000-10-04  
; PRIOR APPLICATION NUMBER: US 60/236,359  
; PRIOR FILING DATE: 2000-09-27  
; PRIOR APPLICATION NUMBER: PCT/US01/00666  
; PRIOR FILING DATE: 2001-01-30  
; PRIOR APPLICATION NUMBER: PCT/US01/00667  
; PRIOR FILING DATE: 2001-01-30  
; PRIOR APPLICATION NUMBER: PCT/US01/00664  
; PRIOR FILING DATE: 2001-01-30  
; PRIOR APPLICATION NUMBER: PCT/US01/00669  
; PRIOR FILING DATE: 2001-01-30  
; PRIOR APPLICATION NUMBER: PCT/US01/00665  
; PRIOR FILING DATE: 2001-01-30  
; PRIOR APPLICATION NUMBER: PCT/US01/00668  
; PRIOR FILING DATE: 2001-01-30  
; PRIOR APPLICATION NUMBER: PCT/US01/00663  
; PRIOR FILING DATE: 2001-01-30  
; PRIOR APPLICATION NUMBER: PCT/US01/00662  
; PRIOR FILING DATE: 2001-01-30  
; PRIOR APPLICATION NUMBER: PCT/US01/00661  
; PRIOR FILING DATE: 2001-01-30



GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: February 17, 2004, 10:50:13 ; Search time 6.33663 Seconds  
(without alignments)  
148.428 Million cell updates/sec

Title: US-09-900-147-4

Perfect score: 101

Sequence: 1 NVLMAMNIISKEKEIKWIG 20

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 127863 seqs, 47026705 residues

Total number of hits satisfying chosen parameters: 13973

Minimum DB seq length: 0

Maximum DB seq length: 100

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt\_41.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	38	37.6	65	1 V089_FOWPV	072899 fowlpox vir
2	36	35.6	45	1 HLD_STAM	P01506 staphylococ
3	36	35.6	64	1 YC53_ARCFU	O29015 archaeoglob
4	36	35.6	88	1 YDFK_ECOLI	P76154 escherichia
5	36	35.6	88	1 YNAE_ECOLI	P76073 escherichia
6	35	34.7	74	1 Y295_ARCFU	O29947 archaeoglob
7	35	34.7	95	1 Y511_BUCAP	Q8k946 buchnera ap
8	35	34.7	98	1 RS24_THEAC	Q9hj79 thermoplas
9	34.5	34.2	90	1 RL16_STR3	Q8e4h3 streptococc
10	34	33.7	72	1 RL15_BACLI	P35138 bacillus li
11	34	33.7	91	1 VAPD_HABIN	P71351 haemophilus
12	34	33.7	95	1 RS14_CARRU	Q9aif4 carsonella
13	34	33.7	95	1 Y117_NPVAC	P41670 autographa
14	34	33.7	97	1 SPAC_BP14	P39230 bacterioph
15	33.5	33.2	90	1 RS16_STRPY	P58124 streptococc
16	33.5	33.2	91	1 RS16_STRPY	Q8dun9 streptococc
17	33	32.7	42	1 RL32_MAZE	P51421 zea mays (m
18	33	32.7	70	1 CSPR_RICCN	Q92gv1 rickettsia
19	33	32.7	70	1 CSPR_RICPR	Q9zcp9 rickettsia
20	33	32.7	72	1 VB11_VACCV	Q01229 vaccinia vi
21	33	32.7	88	1 RL29_SULTO	Q97518 sulfolobus
22	33	32.7	88	1 VB11_VACCC	P21007 vaccinia vi
23	32	31.7	35	1 GBGU_MOUSE	Q61017 mus musculu
24	32	31.7	56	1 Y546_METJA	O57966 methanococc
25	32	31.7	59	1 SECE_BACLI	P38381 bacillus li
26	32	31.7	59	1 SECE_BACSU	Q06799 bacillus su
27	32	31.7	76	1 DMS4_PHYBI	P81486 phylomedus
28	32	31.7	84	1 LEU2_HUMAN	O43262 homo sapien
29	32	31.7	86	1 YC15_METJA	O58612 methanococc
30	32	31.7	87	1 BOPA_BACSU	P24282 bacillus su
31	32	31.7	90	1 RL15_LOTJA	Q9bbn7 lotus japon
32	32	31.7	90	1 RL15_SPIOL	Q9m314 spinacia ol
33	32	31.7	90	1 Y530_BUCAI	P57596 buchnera ap

RESULT 1  
V089\_FOWPV STANDARD; PRT; 65 AA.  
ID V089\_FOWPV

AC 072899;  
DT 16-OCT-2001 (Rel. 40, Created)

DT 16-OCT-2001 (Rel. 40, Last sequence update)

DT 16-OCT-2001 (Rel. 40, Last annotation update)

DE Protein FPV089.

GN FPV089 OR FP12L.

OS Fowlpox virus (FPV).

OC Viruses; dsDNA viruses, no RNA stage; Poxviridae; Chordopoxvirinae;

OC Avipoxvirus

OX NCBI\_TaxID=10261;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=FP-9 / Isolate HP-440;

RX MEDLINE=98451804; PubMed=9778782;

RA Pollitt E., Skinner M.A., Heaphy S.;

RT "Nucleotide sequence of the 4.3 kbp BamHI-N fragment of fowlpox virus

FP9.";

RL Virus Genes 17:5-9 (1998).

RN [2]

RP SEQUENCE FROM N.A.

RX MEDLINE=20193820; PubMed=10729156;

RA Alfonso C.L., Tulman E.R., Lu Z., Zaak L., Kutish G.F., Rock D.L.;

RT "The genome of fowlpox virus.";

RL J. Virol. 74:3815-3831(2000).

CC -I- SIMILARITY: BELONGS TO THE POXVIRUSES I2 FAMILY.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration

CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -

CC the European Bioinformatics Institute. There are no restrictions on its

CC use by non-profit institutions as long as its content is in no way

CC modified and this statement is not removed. Usage by and for commercial

CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>

CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

CC -----

EMBL; AJ223385; CAA11292.1; -

DR EMBL; AF198100; AAF44433.1; -

SQ SEQUENCE 65 AA; 7573 MW; ECDD9A25B824DF92 CRC64;

Query Match 37.6%; Score 38; DB 1; Length 65;

Best Local Similarity 31.8%; Pred. No. 22;

Matches 6; Conservative 5; Mismatches 8; Indels 0; Gaps 0;

QY 1 NVLMAMNIISKEKEIKWI 19

DB 25 NTIMTWTGKESKQLSWL 43

RESULT 2

HLD\_STAM STANDARD; PRT; 45 AA.

ID HLD\_STAM

AC P01506;

DT 21-JUL-1986 (Rel. 01, Created)

DT 01-MAY-1991 (Rel. 18, Last sequence update)

P42550 escherichia  
Q57590 methanococc  
P49502 odontella s  
Q58945 methanococc  
P50154 bos taurus  
Q9ppy4 ureaplasma  
P42035 methanobact  
O26898 methanobact  
P00114 synechococc  
P00115 synechococc  
P00098 rhodocycclus  
Q8tsa4 methanosarc

ALIGNMENTS

DT 28-FEB-2003 (Rel. 41, Last annotation update)  
DE Delta-hemolysin precursor (Delta-toxin).  
GN HLD OR SAV2035 OR SAS065.  
OS Staphylococcus aureus (strain Mu50 / ATCC 700699),  
OS Staphylococcus aureus (strain N315), and  
OS Staphylococcus aureus.  
OC Bacteria; Firmicutes; Bacillales; Staphylococcus.  
OX NCBI\_TaxID=158878, 158879, 1280;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=Mu50 / ATCC 700699, and N315;  
RX MEDLINE=21311952; PubMed=1418146;  
RA Kuroda M., Ohta T., Uchiyama I., Baba T., Yuzawa H., Kobayashi I.,  
RA Kanamori M., Matsumaru H., Maruyama A., Murakami H., Hoshino A.,  
RA Mizutani-Ui Y., Takahashi N.K., Sawano T., Inoue R.-I., Kaito C.,  
RA Sekimizu K., Hirakawa H., Kihara S., Goto S., Yabuzaki J.,  
RA Kanehisa M., Yamashita A., Oshima K., Furuya K., Yoshino C., Shiba T.,  
RA Hattori M., Ogasawara N., Hayashi H., Hiramoto K.,  
RT "Whole genome sequencing of methicillin-resistant Staphylococcus  
RT aureus";  
RL Lancet 357:1225-1240 (2001).  
RN [2]  
RP SEQUENCE FROM N.A.  
RC STRAIN=NCTC 8325-4;  
RX MEDLINE=90158509; PubMed=2622452;  
RA Janson L., Loeffelholz S., Arvidson S.,  
RT "Identification and nucleotide sequence of the delta-lysin gene, hld,  
RT adjacent to the accessory gene regulator (agr) of Staphylococcus  
RT aureus";  
RL Mol. Gen. Genet. 219:480-485 (1989).  
RN [3]  
RP SEQUENCE FROM N.A.  
RC STRAIN=Isolate GAL;  
RX MEDLINE=96004766; PubMed=7565609;  
RA Novick R.P., Projan S.J., Kornblum J., Ross H.F., Ji G.,  
RA Kreiswirth B., Vandenesch F., Mochales S.,  
RT "The agr P2 operon: an autocatalytic sensory transduction system in  
RT Staphylococcus aureus";  
RL Mol. Gen. Genet. 248:446-458 (1995).  
RN [4]  
RP SEQUENCE FROM N.A.  
RC STRAIN=RN4220;  
RA Bischoff M.;  
RT "DNA for hld, agrB, and agrD genes of Staphylococcus aureus strain  
RT RN4220";  
RL Submitted (FEB-2000) to the EMBL/GenBank/DBJ databases.  
RN [5]  
RP SEQUENCE OF 20-45.  
RC STRAIN=186X;  
RX MEDLINE=80246463; PubMed=7398877;  
RA Fitton J.E., Dell A., Shaw W.V.;  
RT "The amino acid sequence of the delta haemolysin of Staphylococcus  
RT aureus";  
RL FEBS Lett. 115:209-212 (1980).  
RN [6]  
RP SEQUENCE OF 20-45.  
RC STRAIN=Canine variant;  
RA Fitton J.E., Hunt D.F., Marasco J., Shabanowitz J., Winston S.,  
RA Dell A.;  
RT "The amino acid sequence of delta haemolysin purified from a canine  
RT isolate of S. aureus";  
RL FEBS Lett. 169:25-29 (1984).  
RN [7]  
RP 3D-STRUCTURE MODELING.  
RX MEDLINE=91126037; PubMed=2281085;  
RA Raghunathan G., Seetharamulu P., Brooks B.R., Guy H.R.;  
RT "Models of delta-hemolysin membrane channels and crystal structures";  
RL Proteins 8:213-225 (1990).  
CC -1- FUNCTION: DELTA-TOXIN Lyses erythrocytes and many other mammalian  
CC cells.  
CC -1- SUBCELLULAR LOCATION: ASSUMED TO INSERT AND OLIGOMERIZE IN  
CC MEMBRANES.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
DR EMBL; AP003364; BAB58197.1; ALT\_INIT.  
DR EMBL; AP003365; BAB43122.1; ALT\_INIT.  
DR EMBL; X52543; CAA36780.1; ALT\_INIT.  
DR EMBL; AF230358; AAF43204.1; --  
DR PIR; A01767; LESAD.  
DR PIR; A89995; A89995.  
DR PDB; 1DHL; 15-OCT-94.  
DR PDB; 2DTB; 31-OCT-93.  
DR PDB; 1DTC; 31-OCT-93.  
DR Pfam; PF05372; Delta\_lysin; 1.  
KW Hemolysis; Formylation; Toxin; Transmembrane; 3D-structure;  
KW Complete proteome.  
FT PROPEP 1 19  
FT CHAIN 20 45  
FT MOD RES 20 20  
FT VARIANT 22 22  
FT VARIANT 29 31  
FT VARIANT 34 34  
FT VARIANT 36 37  
FT VARIANT 40 40  
FT VARIANT 43 43  
FT HELIX 21 41  
SQ SEQUENCE 45 AA; 5140 MW; 3B6B661E0342CA01 CRC64;  
Query Match 35.6%; Score 36; DB 1; Length 45;  
Best Local Similarity 44.4%; Pred. No. 32;  
Matches 8; Conservative 3; Mismatches 7; Indels 0; Gaps 0;  
QY 2 VLMAMNIISKEKEIKWI 19  
DB 18 ISMAQDIISTIGDLVKWI 35  
RESULT 3  
YCS3\_ARCFU  
ID YCS3\_ARCFU STANDARD; PRT; 64 AA.  
AC Q29015;  
DT 16-OCT-2001 (Rel. 40, Created)  
DT 16-OCT-2001 (Rel. 40, Last sequence update)  
DT 16-OCT-2001 (Rel. 40, Last annotation update)  
DE Hypothetical protein AF1253.  
GN AF1253.  
OS Archaeoglobus fulgidus.  
OC Archaea; Euryarchaeota; Archaeoglobi; Archaeoglobales;  
OC Archaeoglobaceae; Archaeoglobus.  
OX NCBI\_TaxID=2234;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=VC-16 / DSM 4304 / ATCC 49558;  
RX MEDLINE=98049343; PubMed=9389475;  
RA Klenk H.-P., Clayton R.A., Tomb J.-F., White O., Nelson K.E.,  
RA Ketchum K.A., Dodson R.J., Gwinn M., Hickey E.K., Peterson J.D.,  
RA Richardson D.L., Kerlavage A.R., Graham D.E., Kyrpides N.C.,  
RA Fleischmann R.D., Quackenbush J., Lee N.H., Sutton G.G., Gill S.,  
RA Kirkness E.F., Dougherty B.A., McKenney K., Adams M.D., Loftus B.,  
RA Peterson S., Reich C.I., McNeil L.K., Badger J.H., Glodek A., Zhou L.,  
RA Overbeek R., Gocayne J.D., Weidman J.F., McDonald L., Utterback T.,  
RA Cotton M.D., Spriggs T., Artiach P., Kaine B.P., Sykes S.M.,  
RA Sadow P.W., D'Andrea K.P., Bowman C., Fujii C., Garland S.A.,  
RA Mason T.M., Olsen G.J., Fraser C.M., Smith H.O., Woese C.R.,  
RA Venter J.C.;  
RT "The complete genome sequence of the hyperthermophilic, sulphate-  
RT reducing archaeon Archaeoglobus fulgidus";  
RL Nature 390:364-370 (1997).

```

CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; AE001018; AAB90006.1; -.
CC DR PIR; D69406; D69406.
CC DR TIGR; AF1253; -.
CC KW Hypothetical protein; Complete proteome.
CC SQ SEQUENCE 64 AA; 7418 MW; CD6CF9F885F4FF15 CRC64;

Query Match 35.6%; Score 36; DB 1; Length 64;
Best Local Similarity 60.0%; Pred. No. 46;
Matches 9; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

QY 5 ANMIISKEKKEIKW 19
Db |:|:|:|:|:|:|
47 AVEEITKAKKEIKKI 61

RESULT 4
YDFK_ECOLI
ID YDFK_ECOLI STANDARD; PRT; 88 AA.
AC P76154;
DT 15-JUL-1998 (Rel. 36, Created)
DT 15-JUL-1998 (Rel. 36, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE Hypothetical protein ydfk.
GN YDFK OR B1544.
OS Escherichia coli.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Escherichia.
OX NCBI_TaxID=562;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=K12 / MG1655;
RX MEDLINE=97426617; PubMed=9278503;
RA Blattner F.R., Plunkett G. III, Bloch C.A., Perna N.T., Burland V.,
RA Riley M., Collado-Vides J., Glaesner J.D., Rode C.K., Mayhew G.F.,
RA Gregor J., Davis N.W., Kirkpatrick H.A., Goeden M.A., Rose D.J.,
RA Mau B., Shao Y.;
RT "The complete genome sequence of Escherichia coli K-12.";
RL Science 277:1453-1474(1997).
CC -!- SIMILARITY: STRONG, TO E.COLI YDFK.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; AE000252; AAC74617.1; -.
CC DR PIR; C64909; C64909.
CC DR EcoGene; EG13823; ydfk.
CC KW Hypothetical protein; Complete proteome.
CC SQ SEQUENCE 88 AA; 10136 MW; 534B06E2287CC6A0 CRC64;

Query Match 35.6%; Score 36; DB 1; Length 88;
Best Local Similarity 44.4%; Pred. No. 63;
Matches 8; Conservative 2; Mismatches 6; Indels 2; Gaps 1;

QY 1 NVLMAMNIIISKEKKEIKW 18
Db |:|:|:|:|:|:|
5 NIL--RNIFMKSKDTLKW 20

RESULT 5
YDFK_ECOLI
ID YDFK_ECOLI STANDARD; PRT; 88 AA.
AC P76154;
DT 15-JUL-1998 (Rel. 36, Created)
DT 15-JUL-1998 (Rel. 36, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE Hypothetical protein ydfk.
GN YDFK OR B1544.
OS Escherichia coli.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Escherichia.
OX NCBI_TaxID=562;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=K12 / MG1655;
RX MEDLINE=97426617; PubMed=9278503;
RA Blattner F.R., Plunkett G. III, Bloch C.A., Perna N.T., Burland V.,
RA Riley M., Collado-Vides J., Glaesner J.D., Rode C.K., Mayhew G.F.,
RA Gregor J., Davis N.W., Kirkpatrick H.A., Goeden M.A., Rose D.J.,
RA Mau B., Shao Y.;
RT "The complete genome sequence of Escherichia coli K-12.";
RL Science 277:1453-1474(1997).
CC -!- SIMILARITY: STRONG, TO E.COLI YNAE.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; AE000234; AAC74457.1; -.
CC DR PIR; B64888; B64888.
CC DR EcoGene; EG13373; ynaE.
CC KW Hypothetical protein; Complete proteome.
CC SQ SEQUENCE 88 AA; 10109 MW; 534B1813C77CC6A0 CRC64;

Query Match 35.6%; Score 36; DB 1; Length 88;
Best Local Similarity 44.4%; Pred. No. 63;
Matches 8; Conservative 2; Mismatches 6; Indels 2; Gaps 1;

QY 1 NVLMAMNIIISKEKKEIKW 18
Db |:|:|:|:|:|:|
5 NIL--RNIFMKSKDTLKW 20

RESULT 6
Y295_ARCFU
ID Y295_ARCFU STANDARD; PRT; 74 AA.
AC Q29947;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Hypothetical protein AF0295.
GN AF0295.
OS Archaeoglobus fulgidus.
OC Archaea; Euryarchaeota; Archaeoglobi; Archaeoglobales;
OC Archaeoglobaceae; Archaeoglobus.
OX NCBI_TaxID=2234;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=VC-16 / DSM 4304 / ATCC 49558;
RX MEDLINE=98049343; PubMed=9389475;
RA Klenk H.-P., Clayton R.A., Tomb J.-F., White O., Nelson K.E.,
RA Ketchum K.A., Dodson R.J., Gwinn M., Hickey E.K., Peterson J.D.,
RA Richardson D.L., Kerlavage A.R., Graham D.E., Kyrpides N.C.,
RA Fleischmann R.D., Quackenbush J., Lee N.H., Sutton G.G., Gill S.,
RA Kirkness E.F., Dougherty B.A., McKenney K., Adams M.D., Loftus B.,
RA Peterson S., Reich C.I., McNeil L.K., Badger J.H., Glodek A., Zhou L.,
RA Overbeek R., Gocayne J.D., Weidman J.K., Badger J.H., Glodek A., Zhou L.,
RA Cotton M.D., Spriggs T., Artiach P., Kaine B.P., Sykes S.M.,
RA Sadow P.W., D'Andrea K.P., Bowman C., Fujii C., Garland S.A.,

```

```

YNAE_ECOLI
ID YNAE_ECOLI STANDARD; PRT; 88 AA.
AC P76073;
DT 15-JUL-1998 (Rel. 36, Created)
DT 15-JUL-1998 (Rel. 36, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE Hypothetical protein ynaE.
GN YNAE OR B1375.
OS Escherichia coli.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Escherichia.
OX NCBI_TaxID=562;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=K12 / MG1655;
RX MEDLINE=97426617; PubMed=9278503;
RA Blattner F.R., Plunkett G. III, Bloch C.A., Perna N.T., Burland V.,
RA Riley M., Collado-Vides J., Glaesner J.D., Rode C.K., Mayhew G.F.,
RA Gregor J., Davis N.W., Kirkpatrick H.A., Goeden M.A., Rose D.J.,
RA Mau B., Shao Y.;
RT "The complete genome sequence of Escherichia coli K-12.";
RL Science 277:1453-1474(1997).
CC -!- SIMILARITY: STRONG, TO E.COLI YDFK.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; AE000234; AAC74457.1; -.
CC DR PIR; B64888; B64888.
CC DR EcoGene; EG13373; ynaE.
CC KW Hypothetical protein; Complete proteome.
CC SQ SEQUENCE 88 AA; 10109 MW; 534B1813C77CC6A0 CRC64;

Query Match 35.6%; Score 36; DB 1; Length 88;
Best Local Similarity 44.4%; Pred. No. 63;
Matches 8; Conservative 2; Mismatches 6; Indels 2; Gaps 1;

QY 1 NVLMAMNIIISKEKKEIKW 18
Db |:|:|:|:|:|:|
5 NIL--RNIFMKSKDTLKW 20

RESULT 6
Y295_ARCFU
ID Y295_ARCFU STANDARD; PRT; 74 AA.
AC Q29947;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Hypothetical protein AF0295.
GN AF0295.
OS Archaeoglobus fulgidus.
OC Archaea; Euryarchaeota; Archaeoglobi; Archaeoglobales;
OC Archaeoglobaceae; Archaeoglobus.
OX NCBI_TaxID=2234;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=VC-16 / DSM 4304 / ATCC 49558;
RX MEDLINE=98049343; PubMed=9389475;
RA Klenk H.-P., Clayton R.A., Tomb J.-F., White O., Nelson K.E.,
RA Ketchum K.A., Dodson R.J., Gwinn M., Hickey E.K., Peterson J.D.,
RA Richardson D.L., Kerlavage A.R., Graham D.E., Kyrpides N.C.,
RA Fleischmann R.D., Quackenbush J., Lee N.H., Sutton G.G., Gill S.,
RA Kirkness E.F., Dougherty B.A., McKenney K., Adams M.D., Loftus B.,
RA Peterson S., Reich C.I., McNeil L.K., Badger J.H., Glodek A., Zhou L.,
RA Overbeek R., Gocayne J.D., Weidman J.K., Badger J.H., Glodek A., Zhou L.,
RA Cotton M.D., Spriggs T., Artiach P., Kaine B.P., Sykes S.M.,
RA Sadow P.W., D'Andrea K.P., Bowman C., Fujii C., Garland S.A.,

```

```

RA Mason T.M., Olsen G.J., Fraser C.M., Smith H.O., Woese C.R.,
RA Venter J.C.;
RT "The complete genome sequence of the hyperthermophilic, sulphate-
RT reducing archaeon Archaeoglobus fulgidus.";
RL Nature 390:364-370 (1997).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; AE001084; AAB90939.1; -.
DR PIR; G69286; G69286.
DR TIGR; AF0295; -.
DR KW Hypothetical protein; Complete proteome.
SQ SEQUENCE 74 AA; 8839 MW; 0DCE58F3B44B2F5 CRC64;
Query Match 34.7%; Score 35; DB 1; Length 74;
Best Local Similarity 52.9%; Pred. No. 77;
Matches 9; Conservative 1; Mismatches 7; Indels 0; Gaps 0;
QY 2 VLMAMNIIISKEIKW 18
DB 5 VLEAGKIISPNEKVIW 21
-----
RESULT 7
Y511_BUCAP STANDARD; PRT; 95 AA.
ID Y511_BUCAP
AC Q8K946;
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Hypothetical protein BUG511.
GN BUG511.
OS Buchnera aphidicola (subsp. Schizaphis graminum).
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Buchnera.
OX NCBI_TaxID=98794;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=22084549; PubMed=12089438;
RA Tamás I., Klasson L., Canbaeck B., Naeslund A.K., Eriksson A.-S.,
RA Wernegreen J.J., Sandstroem J.P., Moran N.A., Andersson S.G.E.;
RT "50 million years of genomic stasis in endosymbiotic bacteria.";
RL Science 296:2376-2379 (2002).
CC -!- SIMILARITY: STRONG, TO E. COLI YHEI.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; AE014125; AAM68054.1; -.
DR PIR; PF04077; DerH; 1.
DR KW Hypothetical protein; Complete proteome.
SQ SEQUENCE 95 AA; 10992 MW; 51E8A2023BE57141 CRC64;
Query Match 34.7%; Score 35; DB 1; Length 95;
Best Local Similarity 46.2%; Pred. No. 99;
Matches 6; Conservative 3; Mismatches 4; Indels 0; Gaps 0;
QY 6 MNIISKEKKEIKW 18
DB 83 VNLTKNKQIIW 95
-----

```

```

RESULT 8
RS24_THEAC STANDARD; PRT; 98 AA.
ID RS24_THEAC
AC Q9HJ79;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE 30S ribosomal protein S24e.
GN RPS2E OR TAI092.
OS Thermoplasma acidophilum.
OC Archaea; Euryarchaeota; Thermoplasmata; Thermoplasmatales;
OC Thermoplasmataceae; Thermoplasma.
OX NCBI_TaxID=2303;
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN=DSM 1728;
RX MEDLINE=20479972; PubMed=11029001;
RA Ruepp A., Graml W., Santos-Martinez M.-L., Koretke K.K., Volker C.,
RA Mewes H.-W., Frishman D., Stocker S., Lupas A.N., Baumeister W.;
RT "The genome sequence of the thermoacidophilic scavenger Thermoplasma
RT acidophilum."
RL Nature 407:508-513 (2000).
CC -!- SIMILARITY: BELONGS TO THE S24E FAMILY OF RIBOSOMAL PROTEINS.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; AL445066; CAC12220.1; -.
DR HAMAP; MF_00545; -.
DR InterPro; IPR001976; Ribosomal_S24E.
DR Pfam; PF01282; Ribosomal_S24e; 1.
DR ProDom; PD006052; Ribosomal_S24e; 1.
DR PROSITE; PS00529; RIBOSOMAL_S24E; 1.
KW Ribosomal protein; Complete proteome.
SQ SEQUENCE 98 AA; 11516 MW; 036ED080443CF8C1 CRC64;
Query Match 34.7%; Score 35; DB 1;
Best Local Similarity 40.9%; Pred. No. 1e+02;
Matches 9; Conservative 4; Mismatches 1; Indels 8; Gaps 1;
QY 6 MNIISKEK-----KEIKWI 19
DB 1 MDLIIEKEDNPILKKEIKYV 22
-----
RESULT 9
RS16_STR33 STANDARD; PRT; 90 AA.
ID RS16_STR33
AC Q8E4H3; Q8DYW5;
DT 15-SEP-2003 (Rel. 42, Created)
DT 15-SEP-2003 (Rel. 42, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE 30S ribosomal protein S16.
GN RPS16 OR GBS1428 OR SAG1358.
OS Streptococcus agalactiae (serotype III), and
OS Streptococcus agalactiae (serotype V).
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC Streptococcus.
OX NCBI_TaxID=216495, 216466;
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN=NEW316 / Serotype III;
RX MEDLINE=22242508; PubMed=12354221;
RA Glaser P., Rusniok C., Buchrieser C., Chevallier F., Frangeul L.,
RA Meadek T., Zouine M., Couve E., Lailouli L., Foyart C., Trieu-Cuot P.,
RA Kunst F.;
RT "Genome sequence of Streptococcus agalactiae, a pathogen causing
RT invasive neonatal disease."

```

```

RL Mol. Microbiol. 45:1499-1513(2002).
RN [2]
RP SEQUENCE FROM N.A.
RX STRAIN=2603 V/R / Serotype V;
RX MEDLINE=22222988; PubMed=12200547;
RA Tettelin H., Masignani V., Cieslewicz M.J., Eisen J.A., Peterson S.,
RA Wessels M.R., Paulsen I.T., Nelson K.E., Margarit I., Read T.D.,
RA Madoff L.C., Wolf A.M., Beanan M.J., Brinkac L.M., Daugherty S.C.,
RA DeBoy R.T., Durkin A.S., Kolonay J.F., Madupu R., Lewis M.R.,
RA Radune D., Fedorova N.B., Scanlan D., Khouri H., Mulligan S.,
RA Carby H.A., Cline R.T., Van Aken S.E., Gill J., Scarselli M., Mora M.,
RA Iacobini E.T., Brettoni C., Galli G., Mariani M., Vegni F., Maione D.,
RA Rinaudo D., Rappuoli R., Telford J.B., Kasper D.B., Grandi G.,
RA Fraser C.M.;
RT "Complete genome sequence and comparative genomic analysis of an
RT emerging human pathogen, serotype V Streptococcus agalactiae.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:12391-12396(2002).
CC -1- SIMILARITY: BELONGS TO THE S16P FAMILY OF RIBOSOMAL PROTEINS.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; AL766850; CAD47087.1; -
DR EMBL; AE014254; AAN00229.1; -
DR Sagalists; gbs1428; -
DR TIGR; SAG1358; -
DR HAMAP; MF_00385; -; 1.
DR InterPro; IPR000307; Ribosomal_S16.
DR Pfam; PF00886; Ribosomal_S16; 1.
DR ProDom; PD003791; Ribosomal_S16; 1.
DR TIGRFAMs; TIGR00002; S16; 1.
DR PROSITE; PS00732; RIBOSOMAL_S16; FALSE_NEG.
KW Ribosomal protein; Complete proteome.
SQ SEQUENCE 90 AA; 10282 MW; CDCAY63D78007CBE CRC64;

Query Match 34.2%; Score 34.5; DB 1; Length 90;
Best Local Similarity 35.0%; Pred. No. 1.1e+02;
Matches 7; Conservative 8; Mismatches 4; Indels 1; Gaps 1;

QY 1 NVLMANNIIS-KEKKEIKWI 19
Db 41 NPLVAENQVIKERVLEWL 60

RESULT 10
RL15_BACLI
ID RL15_BACLI STANDARD; PRT; 72 AA.
AC P35138;
DT 01-FEB-1994 (Rel. 28, Created)
DT 01-FEB-1994 (Rel. 28, Last sequence update)
DT 01-FEB-1994 (Rel. 28, Last annotation update)
DE 50S ribosomal protein L15 (Fragment).
GN RPL0.
OS Bacillus licheniformis.
OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.
OX NCBI_TaxID=1402;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=93062802; PubMed=1435726;
RA Tschauder S., Driessen A.J.M., Freudl R.;
RT "Cloning and molecular characterization of the secY genes from
RT Bacillus licheniformis and Staphylococcus carnosus: comparative
RT analysis of nine members of the SecY family.";
RL Mol. Genet. 235:147-152(1992).
CC -1- FUNCTION: THIS PROTEIN BINDS DIRECTLY TO 23S RIBOSOMAL RNA.
CC -1- SIMILARITY: BELONGS TO THE L15P FAMILY OF RIBOSOMAL PROTEINS.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration

```

```

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; X70087; -; NOT ANNOTATED_CDS.
DR FIR; S34404; S34404.
DR InterPro; IPR001196; Ribosomal_L15.
DR Pfam; PF00256; L15; 1.
DR Pfam; PF01305; Ribosomal_L15; 1.
DR PROSITE; PS00475; RIBOSOMAL_L15; 1.
KW Ribosomal protein; rRNA-binding.
FT NON TER 1
SQ SEQUENCE 72 AA; 7441 MW; 6DC3BA681492E6C9 CRC64;

Query Match 33.7%; Score 34; DB 1; Length 72;
Best Local Similarity 36.8%; Pred. No. 1.1e+02;
Matches 7; Conservative 5; Mismatches 7; Indels 0; Gaps 0;

QY 2 VLMANNIISKEKEIKWIG 20
Db 21 LLLETGVISKLSGVKILG 39

RESULT 11
VAPD_HAEIN
ID VAPD_HAEIN STANDARD; PRT; 91 AA.
AC P71351;
DT 15-JUL-1999 (Rel. 38, Created)
DT 15-JUL-1999 (Rel. 38, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Virulence-associated protein D.
GN VAPD OR HI0450
OS Haemophilus influenzae.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Pasteurellales;
OC Pasteurellaceae; Haemophilus.
OX NCBI_TaxID=727;
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN=Rd / KW20 / ATCC 51907;
RX MEDLINE=95350630; PubMed=7542800;
RA Fleischmann R.D., Adams M.D., White O., Clayton R.A., Kirkness E.F.,
RA McKenney K., Sutton G., Fitzhugh W., Fields C.A., Gocayne J.D.,
RA Scott J.D., Shirley R., Liu L.-I., Glodek A., Kelley J.M.,
RA Weidman J.F., Phillips C.A., Spriggs T., Hedblom E., Cotton M.D.,
RA Utterback T.R., Hanna M.C., Nguyen D.T., Saudek D.M., Brandon R.C.,
RA Fine L.D., Fritchman J.L., Fuhrmann J.B., Geoghagen N.S.M.,
RA Gnehm C.L., McDonald L.A., Small K.V., Fraser C.M., Smith H.O.,
RA Venter J.C.;
RT "Whole-genome random sequencing and assembly of Haemophilus influenzae
RT Rd.";
RL Science 269:496-512(1995).
RN [2]
RP MEDLINE=20137488; PubMed=10675023;
RA Langen H., Takacs B., Evers S., Berndt P., Lahm H.W., Wipf B.,
RA Gray C., Fountoulakis M.;
RT "Two-dimensional map of the proteome of Haemophilus influenzae.";
RL Electrophoresis 21:411-429(2000).
CC -1- SIMILARITY: BELONGS TO THE VAPD FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; U32728; AAC22108.1; -

```





CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

DR EMBL; S57514; AAB25710.1; -;  
 DR EMBL; AF158101; AAD42510.1; -;  
 DR PIR; C45681; C45681.  
 KW Signal; Periplasmic.  
 FT CHAIN 1 22 POTENTIAL.  
 FT CHAIN 23 97 SPACKLE PROTEIN.  
 SQ SEQUENCE 97 AA; 10994 MW; E0A5E5E076C97965 CRC64;

Query Match 33.7%; Score 34; DB 1; Length 97;  
 Best Local Similarity 41.7%; Pred. No. 1.5e+02;  
 Matches 5; Conservative 5; Mismatches 2; Indels 0; Gaps 0;

QY 6 MNISKEKEIK 17  
 |||: : : : :  
 Db 49 MNIVKDRPEMK 60

RESULT 15  
 RS16 STRPY STANDARD; PRT; 90 AA.  
 AC P58124;  
 DT 16-OCT-2001 (Rel. 40, Created)  
 DT 16-OCT-2001 (Rel. 40, Last sequence update)  
 DT 15-SEP-2003 (Rel. 42, Last annotation update)  
 DE 30S ribosomal protein S16.  
 GN RPS16 OR SPY0840 OR SPY03\_0567 OR SPS1287 OR SPY18\_0900.

OS Streptococcus pyogenes, (serotype M3), and  
 OS Streptococcus pyogenes (serotype M18).  
 OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;  
 OC Streptococcus.  
 OX NCBI\_TaxID=1314, 198466, 186103;  
 RN [1]

SEQUENCE FROM N.A.  
 RC STRAIN=SF370 / ATCC 700294 / Serotype M1;  
 RX MEDLINE=21192684; PubMed=11296296;  
 RA Ferreretti J.J., McShan W.M., Ajdic D.J., Savic D.J., Savic G., Lyon K.,  
 RA Primeaux C., Sezate S., Suvorov A.N., Kenton S., Lai H.S., Lin S.P.,  
 RA Qian Y., Jia H.G., Najjar F.Z., Ren Q., Zhu H., Song L., White J.,  
 RA Yuan X., Clifton S.W., Roe B.A., McLaughlin R.;  
 RT "Complete genome sequence of an M1 strain of Streptococcus pyogenes.";  
 RL Proc. Natl. Acad. Sci. U.S.A. 98:4658-4663(2001).  
 RN [2]

SEQUENCE FROM N.A.  
 RC STRAIN=MGAS315 / Serotype M3;  
 RX MEDLINE=22133808; PubMed=12122206;  
 RA Beres S.B., Sylva G.L., Barbican K.D., Lei B., Hoff J.S.,  
 RA Mammarella N.D., Liu M.-Y., Smoot J.C., Porcella S.F., Parkins L.D.,  
 RA Campbell D.S., Smith T.M., McCormick J.K., Leung D.Y.M.,  
 RA Schlievert P.M., Musser J.M.;  
 RT "Genome sequence of a serotype M3 strain of group A Streptococcus:  
 RT phage-encoded toxins, the high-virulence phenotype, and clone  
 RT emergence.";  
 RL Proc. Natl. Acad. Sci. U.S.A. 99:10078-10083(2002).  
 RN [3]

SEQUENCE FROM N.A.  
 RC STRAIN=SSI-1 / Serotype M3;  
 RA Nakagawa I., Kurokawa K., Nakata M., Tomiyasu Y., Yamashita A.,  
 RA Yamazaki K., Okahashi N., Kawabata S., Yasunaga T., Hattori M.,  
 RA Hayashi H., Hamada S.;  
 RT "The genome of invasive Streptococcus pyogenes; a comparative analysis  
 RT of S. pyogenes SSI-1, SF370 and MGAS8232.";  
 RL Submitted (MAY-2002) to the EMBL/GenBank/DBJ databases.  
 RN [4]

SEQUENCE FROM N.A.  
 RC STRAIN=MGAS8232 / Serotype M18;  
 RX MEDLINE=21927593; PubMed=11917108;  
 RA Smoot J.C., Barbican K.D., Van Gompel J.J., Smoot L.M., Chaussee M.S.,  
 RA Sylva G.L., Sturdevant D.E., Ricklefs S.M., Porcella S.F.,

RA Parkins L.D., Beres S.B., Campbell D.S., Smith T.M., Zhang Q.,  
 RA Kapur V., Daly J.A., Veasy L.G., Musser J.M.;  
 RT "Genome sequence and comparative microarray analysis of serotype M18  
 RT group A Streptococcus strains associated with acute rheumatic fever  
 RT outbreaks.";  
 RL Proc. Natl. Acad. Sci. U.S.A. 99:4668-4673(2002).

CC -!- SIMILARITY: BELONGS TO THE S16P FAMILY OF RIBOSOMAL PROTEINS.  
 CC  
 CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

DR EMBL; AE006534; AAK33770.1; -;  
 DR EMBL; AE014148; AAM79174.1; -;  
 DR EMBL; AF005145; BAC64382.1; -;  
 DR EMBL; AE010019; AAL97551.1; -;  
 DR HSSP; P80379; 1EMW.  
 DR HAMAP; MF 00385; -; 1.  
 DR InterPro; IPR000307; Ribosomal\_S16.  
 DR Pfam; PF00886; Ribosomal\_S16; 1.  
 DR ProDom; PD003791; Ribosomal\_S16; 1.  
 DR TIGRFAMs; TIGR00002; S16; 1.  
 DR PROSITE; PS00732; RIBOSOMAL\_S16; FALSE\_NEG.  
 KW Ribosomal protein; Complete proteome.  
 SQ SEQUENCE 90 AA; 10252 MW; CDC5E92FEAAB660E CRC64;

Query Match 33.2%; Score 33.5; DB 1; Length 90;  
 Best Local Similarity 40.0%; Pred. No. 1.6e+02;  
 Matches 8; Conservative 6; Mismatches 5; Indels 1; Gaps 1;

QY 1 NVLMAMNIIS-KEKKEIKWI 19  
 |||: : : : :  
 Db 41 NPLVAENQITIKEDRVLEWL 60

RESULT 16

RS16 STRMU STANDARD; PRT; 91 AA.  
 ID RS16\_STRMU  
 AC Q8DUN9;  
 DT 15-SEP-2003 (Rel. 42, Created)  
 DT 15-SEP-2003 (Rel. 42, Last sequence update)  
 DT 15-SEP-2003 (Rel. 42, Last annotation update)  
 DE 30S ribosomal protein S16.  
 GN RPS16 OR SMU.865.  
 OS Streptococcus mutans.  
 OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;  
 OC Streptococcus.  
 OX NCBI\_TaxID=1309;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=UAI59 / ATCC 700610 / Serotype C;  
 RX MEDLINE=22295063; PubMed=12397186;  
 RA Ajdic D., McShan W.M., McLaughlin R.E., Savic G., Chang J.,  
 RA Carson M.B., Primeaux C., Tian R., Kenton S., Jia H., Lin S., Qian Y.,  
 RA Li S., Zhu H., Najjar F., Lai H., White J., Roe B.A., Ferreretti J.J.;  
 RT "Genome sequence of Streptococcus mutans UAI59, a cariogenic dental  
 RT pathogen.";  
 RL Proc. Natl. Acad. Sci. U.S.A. 99:14434-14439(2002).  
 CC -!- SIMILARITY: BELONGS TO THE S16P FAMILY OF RIBOSOMAL PROTEINS.  
 CC

CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

DR EMBL; AE014928; AAN58580.1; -;

```
DR HAMAP; MF 00385; -: 1.
DR InterPro; IPR000307; Ribosomal_S16.
DR Pfam; PF00886; Ribosomal_S16; 1.
DR ProDom; PD003791; Ribosomal_S16; 1.
DR TIGRfam; TIGR00002; S16; 1.
DR TRIPSITE; PS00732; RIBOSOMAL_S16; FALSE_NEG.
KW Ribosomal protein; Complete proteome.
SQ SEQUENCE 91 AA; 10410 MW; 1E7329D1BDF6EC2 CRC64;

Query Match 33.2%; Score 33.5; DB 1; Length 91;
Best Local Similarity 33.3%; Pred. No. 1.7e+02;
Matches 7; Conservative 7; Mismatches 6; Indels 1; Gaps 1;

QY 1 NVLMANNIIS-KEKKEIKWIG 20
DB 41 NPLVTENQVTLKEDRILEMIG 61

RESULT 17
RL32_MAIZE STANDARD; PRT; 42 AA.
AC P51421;
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DE 60S ribosomal protein L32 (Fragment).
GN RFL32.
OS Zea mays (Maize).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC PACCAD clade; Panicoideae; Andropogoneae; Zea.
OX NCBI_TaxID=4577;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CV. HD5 X HD7;
RA Bates E.M., Vergne P., Dumas C.;
RA Submitted (NOV-1993) to the EMBL/GenBank/DBJ databases.
CC -1- SIMILARITY: BELONGS TO THE L32E FAMILY OF RIBOSOMAL PROTEINS.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; X75646; CAA53301.1; -.
CC PIR; S38633; S38633.
CC MaizeDB; 61651; -.
CC InterPro; IPR001515; Ribosomal_L32E.
CC Pfam; PF01655; Ribosomal_L32e; 1.
CC ProDom; PD003823; Ribosomal_L32e; 1.
CC TRIPSITE; PS00580; RIBOSOMAL_L32E; PARTIAL.
KW Ribosomal protein.
FT NON TER 1
SQ SEQUENCE 42 AA; 4709 MW; FA9E004E030D1BF4 CRC64;

Query Match 32.7%; Score 33; DB 1; Length 42;
Best Local Similarity 46.2%; Pred. No. 92;
Matches 6; Conservative 5; Mismatches 2; Indels 0; Gaps 0;

QY 4 MAMNIISKEKEI 16
DB 6 IAHNVSTKKKEI 18

RESULT 18
CSPA_RICCN STANDARD; PRT; 70 AA.
AC Q92GV1;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
KW Ribosomal protein.
FT NON TER 1
SQ SEQUENCE 70 AA; 7771 MW; 822E9714229FE03A CRC64;

Query Match 32.7%; Score 33; DB 1; Length 70;
Best Local Similarity 46.7%; Pred. No. 1.5e+02;
Matches 7; Conservative 2; Mismatches 2; Indels 4; Gaps 1;

QY 4 MAMNIISKEKEIKW 18
DB 1 MATNIVGK---VKW 11

RESULT 19
CSPA_RICPR STANDARD; PRT; 70 AA.
AC Q92CF9;
DT 30-MAY-2000 (Rel. 39, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Cold shock-like protein cspa.
GN RICKETTSIA PROWAZEKII.
OS Rickettsia prowazekii.
OC Bacteria; Proteobacteria; Alphaproteobacteria; Rickettsiales;
OC Rickettsiaceae; Rickettsiidae; Rickettsia.
OX NCBI_TaxID=782;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Madrid E;
RA MEDLINE=99039499; PubMed=9823893;
RA Andersson S.G.E., Zomorodipour A., Andersson J.O.,
RA Sacheritz-Ponten T., Alenmark U.C.M., Podowski R.M., Naeslund A.K.,
RA Eriksson A.-S., Winkler H.H., Kurland C.G.;
RA "The genome sequence of Rickettsia prowazekii and the origin of
RA mitochondria."
RL Nature 396:133-140 (1998).
CC -1- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
CC -1- SIMILARITY: BELONGS TO THE COLD-SHOCK DOMAIN (CSD) FAMILY.
```

```

CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; AJ235272; CRA15108.1; -
DR PIR; B71673; B71673.
DR HSP; P41016; 1C90.
DR InterPro; IPR002059; Cold_shock.
DR Pfam; PF00313; CSF; 1.
DR PRINTS; PR00050; COLDSHOCK.
DR ProDom; PD000621; Cold shock; 1.
DR SMART; SM00357; CSP; 1.
DR PROSITE; PS00352; COLD SHOCK; 1.
KW Transcription regulation; DNA-binding; Activator; Complete proteome.
FT DOMAIN 7 67 CSF.
SQ SEQUENCE 70 AA; 7785 MW; F49FBF2FB344903E CRC64;

Query Match 32.7%; Score 33; DB 1; Length 70;
Best Local Similarity 46.7%; Pred. No. 1.5e+02;
Matches 7; Conservative 2; Mismatches 2; Indels 4; Gaps 1;

QY 4 MAMNIIISKEKKIKW 18
|||::|:|
Db 1 MATNIVGK----VKW 11

RESULT 20
VB11_VACCV STANDARD; PRT; 72 AA.
AC Q01229;
DT 01-APR-1993 (Rel. 25, Created)
DT 01-APR-1993 (Rel. 25, Last sequence update)
DT 01-APR-1993 (Rel. 25, Last annotation update)
DE Protein B11.
GN Vaccinia virus (strain WR).
OS Viruses; dsDNA viruses, no RNA stage; Poxviridae; Chordopoxvirinae;
OC Orthopoxvirus.
OX NCBI_TaxID=10254;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=91259063; PubMed=2045793;
RA Smith G.L., Chan Y.S., Howard S.T.;
RT "Nucleotide sequence of 42 kbp of vaccinia virus strain WR from near
RL J. Gen. Virol. 72:1349-1376(1991).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; D11079; BAA01841.1; -
DR PIR; JQ1805; JQ1805.
SQ SEQUENCE 72 AA; 8184 MW; 372BE3DCE6AFAE07 CRC64;

Query Match 32.7%; Score 33; DB 1; Length 72;
Best Local Similarity 50.0%; Pred. No. 1.6e+02;
Matches 8; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

QY 1 NVLMAMNIIISKEKKEI 16
|||:|:|:|
Db 7 NVEDLINEIDREKKEI 22

```

```

RESULT 21
RL29_SULTO STANDARD; PRT; 88 AA.
AC Q97518;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE 50S ribosomal protein L29P.
GN RPL29P OR STS061.
OS Sulfolobus tokodaii.
OC Archaea; Crenarchaeota; Thermoprotei; Sulfolobales; Sulfolobaceae;
OC Sulfolobus.
OX NCBI_TaxID=111955;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=21456156; PubMed=11572479;
RA Kawarabayasi Y., Hino Y., Horikawa H., Jin-no K., Takahashi M.,
RA Sekine M., Baba S.-I., Ankai A., Kosugi H., Hosoyama A., Fukui S.,
RA Nagai Y., Nishijima K., Otsuka R., Nakazawa H., Takamiya M., Kato Y.,
RA Yoshizawa T., Tanaka T., Kudoh Y., Yamazaki J., Kushida N., Oguchi A.,
RA Aoki K.-I., Masuda S., Yanagii M., Nishimura M., Yamagishi A.,
RA Oshima T., Kikuchi H.;
RT "Complete genome sequence of an aerobic thermoacidophilic
RT Crenarchaeon, Sulfolobus tokodaii strain 7.";
RL DNA Res. 8:123-140(2001).
CC -!- SIMILARITY: BELONGS TO THE L29P FAMILY OF RIBOSOMAL PROTEINS.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; AF000982; BAB65412.1; -
DR HAMAP; MF_00374; atypical; 1.
DR InterPro; IPR001854; Ribosomal L29.
DR Pfam; PF00831; Ribosomal L29; 1.
DR TIGRFAMs; TIGR00012; L29_2.
DR PROSITE; PS00579; RIBOSOMAL_L29; 1.
KW Ribosomal protein; Complete proteome.
SQ SEQUENCE 88 AA; 10277 MW; 9A5B120EB10382C2 CRC64;

Query Match 32.7%; Score 33; DB 1; Length 88;
Best Local Similarity 45.0%; Pred. No. 1.9e+02;
Matches 9; Conservative 4; Mismatches 3; Indels 4; Gaps 1;

QY 2 VLMAMNIIISKE----KKEIK 17
||:|:|:|:|
Db 68 ILTLLSIIKKEIKKESK 87

RESULT 22
VB11_VACCV STANDARD; PRT; 88 AA.
AC P21007;
DT 01-FEB-1991 (Rel. 17, Created)
DT 01-FEB-1991 (Rel. 17, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Protein B11.
GN B11R.
OS Vaccinia virus (strain Copenhagen).
OC Viruses; dsDNA viruses, no RNA stage; Poxviridae; Chordopoxvirinae;
OC Orthopoxvirus.
OX NCBI_TaxID=10249;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=91021027; PubMed=2219722;
RA Goebel S.J., Johnson G.P., Perkus M.E., Davis S.W., Winslow J.P.,
RA Paoletti E.;
RT "The complete DNA sequence of vaccinia virus.";

```

RESULT 25  
SECE BACLI

ID SECE\_BACULI STANDARD; PRT; 59 AA.  
AC P38381;  
DT 01-OCT-1994 (Rel. 30, Created)  
DT 01-OCT-1994 (Rel. 30, Last sequence update)  
DT 01-OCT-1994 (Rel. 30, Last annotation update)  
DE Preprotein translocase secE subunit.  
GN SECE.  
OS Bacillus licheniformis.  
OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.  
OX NCBI\_TaxID=1402;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=88139159; PubMed=3277943;  
RA Dubnau E., Weir J., Nair G., Carter L. III, Moran C.P. Jr., Smith I.;  
RT "Bacillus sporulation gene spoOH codes for sigma 30 (sigma H).";  
RL J. Bacteriol. 170:1054-1062(1988).  
CC -1- FUNCTION: ESSENTIAL FOR PROTEIN EXPORT.  
CC -1- SUBCELLULAR LOCATION: Tail-anchored membrane protein (Potential).  
CC -1- SIMILARITY: Belongs to the secE/SEC61-gamma family.  
CC  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (see <http://www.isb-sib.ch/announcement/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC  
CC EMBL; M29694; -; NOT ANNOTATED\_CDS.  
DR InterPro; IPR001901; SecE.  
DR InterPro; IPR005807; SecE\_bac.  
DR Pfam; PF00584; SecE; 1.  
DR TIGRFAMs; TIGR00964; 3a0501s06; 1.  
DR PROSITE; PS01067; SECE SEC61G; 1.  
KW Protein transport; Translocation; Transmembrane.  
FT TRANSMEM 30 50 POTENTIAL.  
SQ SEQUENCE 59 AA; 6775 MW; BD40479D9FA5837B CRC64;  
  
Query Match 31.7%; Score 32; DB 1; Length 59;  
Best Local Similarity 23.5%; Pred. No. 1.9e+02;  
Matches 4; Conservative 6; Mismatches 7; Indels 0; Gaps 0;  
  
QY 2 VLMAMNIIISKEKEIKW 18  
: : : ||| : :  
Db 3 IIRFLKNVGNKMKVTV 19  
  
RESULT 26  
SECE\_BACSU STANDARD; PRT; 59 AA.  
AC Q06799; P36689;  
DT 01-JUN-1994 (Rel. 29, Created)  
DT 01-JUN-1994 (Rel. 29, Last sequence update)  
DT 28-FEB-2003 (Rel. 41, Last annotation update)  
DE Preprotein translocase secE subunit.  
GN SECE.  
OS Bacillus subtilis.  
OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.  
OX NCBI\_TaxID=1423;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=95058172; PubMed=7968510;  
RA Jeong S., Yoshikawa H., Takahashi H.;  
RT "Isolation and characterization of the secE homologue gene of  
RT Bacillus subtilis.";  
RL Mol. Microbiol. 10:133-142(1993).  
RN [2]  
RP SEQUENCE FROM N.A.  
RC STRAIN=168;  
RX MEDLINE=98044033; PubMed=9384377;  
RA Kunst F., Ogasawara N., Moszer I., Albertini A.M., Alloni G.,  
RA Azevedo V., Bertero M.G., Bessieres P., Bolotin A., Borchert S.,  
RA Borriss R., Boursier L., Brans A., Braun M., Brignell S.C., Bron S.,

RA Brouillet S., Bruschi C.V., Caldwell B., Capuano V., Carter N.M.,  
RA Choi S.K., Codani J.J., Connerton I.F., Cummings N.J., Daniel R.A.,  
RA Denizot F., Devine K.M., Dusterhoft A., Ehrlich S.D., Emmerson P.T.,  
RA Entian K.D., Errington J., Fabret C., Ferrari E., Foulger D.,  
RA Fritz C., Fujita M., Fujita Y., Funa S., Galizzi A., Galleron N.,  
RA Ghim S.Y., Glaser P., Goffeau A., Gollightly E.J., Grandi G.,  
RA Guisepi G., Guy B.J., Haga K., Halech J., Harwood C.R., Henaut A.,  
RA Hilbert H., Holsappel S., Hosono S., Hullo M.F., Itaya M., Jones L.,  
RA Joris B., Karamata D., Kasahara Y., Klaerr-Blanchard M., Klein C.,  
RA Kobayashi Y., Koetter P., Koningsstein G., Krogh S., Kumano M.,  
RA Kurita K., Lapidus A., Lardinois S., Lauber J., Lazarevic V.,  
RA Lee S.M., Levine A., Liu H., Masuda S., Mauel C., Medigue C.,  
RA Medina N., Mellado R.P., Mizuno M., Moestl D., Nakai S., Noback M.,  
RA Noone D., O'Reilly M., Ogawa K., Ogiwara A., Oudega B., Park S.H.,  
RA Parro V., Pohl T.M., Portetelle D., Porwollik S., Prescott A.M.,  
RA Presecan E., Pujic P., Purnelle B., Rapoport G., Rey M., Reynolds S.,  
RA Rieger M., Rivolta C., Rocha E., Roche B., Rose M., Sadaie Y.,  
RA Sato T., Scanlan E., Schleich S., Schroeter R., Scoffone F.,  
RA Sekiguchi J., Sekowska A., Seror S.J., Serror P., Shin B.S., Soldo B.,  
RA Sorokin A., Tacconi E., Takagi T., Takahashi H., Takemaru K.,  
RA Takeuchi M., Tamakoshi A., Tanaka T., Terpetra P., Tognoni A.,  
RA Tosato V., Uchiyama S., Vandenbol M., Vannier F., Vassarotti A.,  
RA Viari A., Wambutt R., Wedler E., Wedler H., Weitzenecker T.,  
RA Winters P., Wipat A., Yamamoto H., Yamane K., Yasumoto K., Yata K.,  
RA Yoshida K., Yoshikawa H.F., Zumstein E., Yoshikawa H., Danchin A.,  
RT "The complete genome sequence of the Gram-positive bacterium Bacillus  
RT subtilis.";  
RL Nature 390:249-256(1997).  
CC -1- FUNCTION: ESSENTIAL FOR PROTEIN EXPORT.  
CC -1- SUBCELLULAR LOCATION: Tail-anchored membrane protein (Potential).  
CC -1- SIMILARITY: Belongs to the secE/SEC61-gamma family.  
CC  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (see <http://www.isb-sib.ch/announcement/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC  
CC EMBL; D13303; BAA02559.1; -;  
DR EMBL; Z99104; CAB11876.1; -;  
DR PIR; S39858; S39858.  
DR Subtilist; BG10161; secE.  
DR InterPro; IPR001901; SecE.  
DR InterPro; IPR005807; SecE\_bac.  
DR Pfam; PF00584; SecE; 1.  
DR TIGRFAMs; TIGR00964; 3a0501s06; 1.  
DR PROSITE; PS01067; SECE SEC61G; 1.  
KW Protein transport; Translocation; Transmembrane; Complete proteome.  
FT TRANSMEM 30 50 POTENTIAL.  
SQ SEQUENCE 59 AA; 6927 MW; F58F9D90EDE4DFC7 CRC64;  
  
Query Match 31.7%; Score 32; DB 1; Length 59;  
Best Local Similarity 40.0%; Pred. No. 1.9e+02;  
Matches 4; Conservative 3; Mismatches 3; Indels 0; Gaps 0;  
  
QY 9 ISKEKEKEIKW 18  
: ||| : : :  
Db 10 VGKEMKKVSW 19  
  
RESULT 27  
DMS4\_PHYBI STANDARD; PRT; 76 AA.  
ID DMS4\_PHYBI  
AC P81486;  
DT 16-OCT-2001 (Rel. 40, Created)  
DT 16-OCT-2001 (Rel. 40, Last sequence update)  
DT 15-SEP-2003 (Rel. 42, Last annotation update)  
DE Dermaseptin BIV precursor (Dermaseptin B4).  
OS Phylomedusa bicolor (two-colored leaf frog).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Amphibia; Batrachia; Anura; Neobatrachia; Bufonoidae; Hylidae;

OC Phylomedusinae; Phylomedusa.  
OX NCBI\_TaxID=8393;  
RN [1]  
RP SEQUENCE FROM N.A., SEQUENCE OF 46-73, AND MASS SPECTROMETRY.  
RC TISSUE=Skin secretion;  
RX MEDLINE=98278974; PubMed=9614066;  
RA Charpentier S., Amiche M., Mester J., Vouille V., Le Caer J.-P.,  
RA Nicolas P., Delfour A., and molecular cloning of dermaseptins B, a  
RT "structure, synthesis, and molecular cloning of dermaseptins B, a  
family of skin peptide antibiotics";  
RL J. Biol. Chem. 273:14690-14697(1998).  
CC -I- FUNCTION: POSSESSES A POTENT ANTIMICROBIAL ACTIVITY AGAINST GRAM-  
CC POSITIVE AND GRAM-NEGATIVE BACTERIA. PROBABLY ACTS BY DISTURBING  
CC MEMBRANE FUNCTIONS WITH ITS AMPHIPATHIC STRUCTURE.  
CC -I- SUBCELLULAR LOCATION: Secreted.  
CC -I- TISSUE SPECIFICITY: Skin.  
CC -I- MASS SPECTROMETRY: MW=2997.15; MW ERR=0.1; METHOD=Electrospray.  
CC -I- SIMILARITY: Belongs to the frog skin active peptide (FSAP) family.  
CC Dermaseptin subfamily.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
DR EMBL; Y16565; CAA76289.1; -;  
DR InterPro; IPR004275; Brevenin.  
DR Pfam; PF03032; Brevenin; 1.  
KW Amphibian defense peptide; Antibiotic; Multigene family; Signal;  
KW Cleavage on pair of basic residues; Amidation.  
FT SIGNAL 1 22 POTENTIAL.  
FT PROPEP 23 43  
FT CHAIN 46 73 DERMASEPTIN BIV.  
FT PROPEP 75 76 POTENTIAL.  
FT MOD RES 73 73 AMIDATION (G-74 PROVIDE AMIDE GROUP).  
SQ SEQUENCE 76 AA; 8642 MW; A8A0525F0709F447 CRC64;  
Query Match 31.7%; Score 32; DB 1; Length 76;  
Best Local Similarity 40.0%; Pred. No. 2.4e+02;  
Matches 6; Conservative 6; Mismatches 3; Indels 0; Gaps 0;  
QY 3 LMANNIISKKEK 17  
Db 16 LVLSICEEEKRENK 30  
RESULT 28  
LEU2\_HUMAN  
ID LEU2\_HUMAN STANDARD; PRT; 84 AA.  
AC O43262;  
DT 16-OCT-2001 (Rel. 40, Created)  
DT 16-OCT-2001 (Rel. 40, Last sequence update)  
DT 28-FEB-2003 (Rel. 41, Last annotation update)  
DE Leukemia associated protein 2 (Deleted in lymphocytic leukemia 2).  
GN LEU2 OR LEU2.  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
OX NCBI\_TaxID=9606;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=98055620; PubMed=9395242;  
RA Liu Y., Corcoran M., Rabool O., Ivanova G., Ibbotson R., Grander D.,  
RA Iyengar A., Baranova A., Kashuba V., Merup M., Wu X., Gardiner A.,  
RA Mullenbach R., Poltarauk A., Hultstrom A.L., Juliusson G., Chapman R.,  
RA Tiller M., Cotter F., Gahrton G., Yankovsky N., Zabarovsky E.,  
RA Einhorn S., Oscier D.;  
RT "Cloning of two candidate tumor suppressor genes within a 10 kb region  
on chromosome 13q14, frequently deleted in chronic lymphocytic  
leukemia.";

RL Oncogene 15:2463-2473(1997).  
CC -I- FUNCTION: MAY ACT AS A TUMOR SUPPRESSOR.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
DR EMBL; Y15228; CAA75516.1; -;  
DR Genew; HGNC:13748; DLEU2.  
DR MIM; 605766; -;  
DR GO; GO:0008181; F:tumor suppressor; TAS.  
KW Anti-oncogene.  
SQ SEQUENCE 84 AA; 10196 MW; D1E021072A0E03B CRC64;  
Query Match 31.7%; Score 32; DB 1; Length 84;  
Best Local Similarity 60.0%; Pred. No. 2.7e+02;  
Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;  
QY 5 ANNIISKKK 14  
Db 73 SLNTIKKKKK 82  
RESULT 29  
YC15\_METJA  
ID YC15\_METJA STANDARD; PRT; 86 AA.  
AC Q58612;  
DT 01-NOV-1997 (Rel. 35, Created)  
DT 01-NOV-1997 (Rel. 35, Last sequence update)  
DT 16-OCT-2001 (Rel. 40, Last annotation update)  
DE Hypothetical protein MJ1215.  
GN MJ1215.  
OS Methanococcus jannaschii.  
OC Archaea; Euryarchaeota; Methanococci; Methanococcales;  
OC Methanocaldococcaceae; Methanocaldococcus.  
OX NCBI\_TaxID=2190;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=JAL-1 / DSM 2661 / ATCC 43067;  
RX MEDLINE=96337999; PubMed=8688087;  
RA Bult C.J., White O., Olsen G.J., Zhou L., Fleischmann R.D.,  
RA Sutton G.G., Blake J.A., Fitzgerald L.M., Clayton R.A., Gocayne J.D.,  
RA Kerlavage A.R., Dougherty B.A., Tomb J.-F., Adams M.D., Reich C.I.,  
RA Overbeek R., Kirkness E.F., Weinstock K.G., Merrick J.M., Glodek A.,  
RA Scott J.L., Geoghagen N.S.M., Weidman J.F., Fuhrmann J.L., Nguyen D.,  
RA Utterback T.R., Kelley J.M., Peterson J.D., Sadow P.W., Hanna M.C.,  
RA Cotton M.D., Roberts K.M., Hurst M.A., Kaine B.P., Borodovsky M.,  
RA Klenk H.-P., Fraser C.M., Smith H.O., Woese C.R., Venter J.C.;  
RT "Complete genome sequence of the methanogenic archaeon, Methanococcus  
jannaschii";  
RL Science 273:1058-1073(1996).  
CC -I- SIMILARITY: BELONGS TO THE M.JANNASCHII MJ0126 / MJ0128 / MJ0141 /  
CC MJ0435 / MJ0604 / MJ1215 / MJ1217 / MJ1305 / MJ1379 FAMILY.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
DR EMBL; U67562; AAB99216.1; -;  
DR PIR; F64451; F64451.  
DR TIGR; MJ1215; -;  
DR InterPro; IPR002934; NTP\_transf.  
DR Pfam; PF01909; NTP\_transf 2; 1.  
KW Hypothetical protein; Complete proteome.  
SQ SEQUENCE 86 AA; 10281 MW; CD9296D0B135FD14 CRC64;

Query Match 31.7%; Score 32; DB 1; Length 86;  
 Best Local Similarity 60.0%; Pred. No. 2.7e+02;  
 Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 8 IISKEKEIK 17  
 Db 14 ILKHKKEK 23

RESULT 30  
 BOFA\_BACSU STANDARD; PRT; 87 AA.

AC P24282;  
 DT 01-MAR-1992 (Rel. 21, Created)  
 DT 01-MAR-1992 (Rel. 21, Last sequence update)  
 DT 28-FEB-2003 (Rel. 41, Last annotation update)  
 DE Sigma-K factor processing regulatory protein BOFA (Bypass-of-forespore protein A).  
 DE BOFA.  
 OS Bacillus subtilis.  
 OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.  
 OX NCBI\_TaxID=1423;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=91088245; PubMed=2124672;  
 RA Alonso C., Shirahige K., Ogasawara N.;  
 RT "Molecular cloning, genetic characterization and DNA sequence  
 analysis of the recM region of Bacillus subtilis."  
 RL Nucleic Acids Res. 18:6771-6777(1990).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=168;  
 RX MEDLINE=96051385; PubMed=7584024;  
 RA Ogasawara N., Nakai S., Yoshikawa H.;  
 RT "Systematic sequencing of the 180 kilobase region of the Bacillus  
 subtilis chromosome containing the replication origin."  
 RL DNA Res. 1:1-14(1994).  
 RN [3]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=168;  
 RX MEDLINE=98044033; PubMed=9384377;  
 RA Kunst F., Ogasawara N., Mosser J., Albertini A.M., Alloni G.,  
 Azevedo V., Bertero M.G., Bessieres P., Bolotin A., Borchert S.,  
 Borries R., Boursier L., Brans A., Braun M., Brignell S.C., Bron S.,  
 Brouillet S., Bruschi C.V., Caldwell B., Capuano V., Carter N.M.,  
 Choi S.K., Codani J.J., Connerton I.F., Cummings N.J., Daniel R.A.,  
 Denizot F., Devine K.M., Dusterhoft A., Ehrlich S.D., Emmerson P.T.,  
 Entian K.D., Errington J., Fabret C., Ferrari E., Foulger D.,  
 Fritz C., Fujita M., Fujita Y., Fuma S., Galizzi A., Galleron N.,  
 Ghim S.Y., Glaser P., Goffeau A., Gollightly E.J., Grandi G.,  
 Guiseppe G., Guy B.J., Haga K., Haeft J., Harwood C.R., Henaut A.,  
 Hilbert H., Holsappel S., Hosono S., Hullo M.F., Itaya M., Jones L.,  
 Joris B., Karamata D., Kasahara Y., Kleaer-Blanchard M., Klein C.,  
 Kobayashi Y., Koetter P., Koningsstein G., Krogh S., Kumano M.,  
 Kurita K., Lapidus A., Lardinois S., Lauber J., Lazarevic V.,  
 Lee S.M., Levine A., Liu H., Masuda S., Mauel C., Medigue C.,  
 Medina N., Mellado R.P., Mizuno M., Moestl D., Nakai S., Noback M.,  
 Noone D., O'Reilly M., Ogawa K., Ogiwara A., Oudega B., Park S.H.,  
 Parro V., Pohl T.M., Portetelle D., Porwollik S., Prescott A.M.,  
 Pressac E., Pujic P., Purnelle B., Rapoport G., Rey M., Reynolds S.,  
 Rieger M., Rivolta C., Rocha E., Roche B., Rose M., Sadaie Y.,  
 Sato T., Scanlan E., Schleich S., Schroeter R., Scoffone F.,  
 Sekiguchi J., Sekowska A., Seror S.J., Serron P., Shin B.S., Soldo B.,  
 Sorokin A., Tacconi E., Takagi T., Takahashi H., Takemaru K.,  
 Takeuchi M., Tamakoshi A., Tanaka T., Terpetra P., Tognoni A.,  
 Tosato V., Uchiyama S., Vandenbol M., Vannier F., Vassarotti A.,  
 Viari A., Wambutt R., Wedler E., Wedler H., Weitzenecker T.,  
 Winters P., Wipat A., Yamamoto H., Yamane K., Yasumoto K., Yata K.,  
 Yoshida K., Yoshikawa H.F., Zumein E., Yoshikawa H., Danchin A.;  
 RT "The complete genome sequence of the Gram-positive bacterium Bacillus  
 subtilis."  
 RL Nature 390:249-256(1997).

[4]  
 CHARACTERIZATION.  
 RP MEDLINE=92250411; PubMed=1577688;  
 RX Ricca E., Cutting S.M., Losick R.;  
 RA "Characterization of bofa, a gene involved in intercompartmental  
 regulation of pro-sigma K processing during sporulation in Bacillus  
 subtilis."  
 RL J. Bacteriol. 174:3177-3184(1992).  
 RN [5]  
 RP TOPOLOGY.  
 RX MEDLINE=97286525; PubMed=9141672;  
 RA Varcamonti M., Marasco R., de Felice M., Sacco M.;  
 RT "Membrane topology analysis of the Bacillus subtilis BofA protein  
 involved in pro-sigma K processing."  
 RL Microbiology 143:1053-1058(1997).  
 CC -I- FUNCTION: INVOLVED IN THE MEDIATION OF THE INTERCOMPARTMENTAL  
 COUPLING OF PRO-SIGMA K PROCESSING TO EVENTS IN THE FORESPORE.  
 CC -I- SUBCELLULAR LOCATION: Integral membrane protein.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 the European Bioinformatics Institute. There are no restrictions on its  
 use by non-profit institutions as long as its content is in no way  
 modified and this statement is not removed. Usage by and for commercial  
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL; X17014; CAA34880.1; -;  
 DR EMBL; D36185; BAA05259.1; -;  
 DR EMBL; Z99104; CAB11799.1; -;  
 DR PIR; B41869; B41869.  
 DR Subtilist; BGI0087; bofa.  
 KW Sporulation; Transmembrane; Complete proteome.  
 FT DOMAIN 1 2 EXTRACELLULAR (PROBABLE).  
 FT TRANSMEM 3 23 PROBABLE.  
 FT DOMAIN 24 36 CYTOPLASMIC (PROBABLE).  
 FT TRANSMEM 37 57 PROBABLE.  
 FT DOMAIN 58 87 EXTRACELLULAR (PROBABLE).  
 SQ SEQUENCE 87 AA; 9010 MW; FF1DA14E6826B70E CRC64;

Query Match 31.7%; Score 32; DB 1; Length 87;  
 Best Local Similarity 71.4%; Pred. No. 2.8e+02;  
 Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 14 KEIKWIG 20  
 Db 25 KPLKNIG 31

RESULT 31  
 RR15 LOTJA STANDARD; PRT; 90 AA.

AC Q9BN7; 2003 (Rel. 41, Created)  
 DT 28-FEB-2003 (Rel. 41, Last sequence update)  
 DT 28-FEB-2003 (Rel. 41, Last annotation update)  
 DE Chloroplast 30S ribosomal protein S15.  
 GN RPS15  
 OS Lotus japonicus.  
 OG Chloroplast.  
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;  
 OC eurosids I; Fabales; Fabaceae; Papilionoideae; Loteae; Lotus.  
 OX NCBI\_TaxID=34305;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=Accession MG-20;  
 RX MEDLINE=21082929; PubMed=11214967;  
 RA Kato T., Kaneko T., Sato S., Nakamura Y., Tabata S.;  
 RT "Complete structure of the chloroplast genome of a legume, Lotus  
 japonicus."  
 RL DNA Res. 7:323-330(2000).  
 CC -I- SIMILARITY: BELONGS TO THE S15P FAMILY OF RIBOSOMAL PROTEINS.



CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 CC EMBL; AF002983; BAB33252.1; --  
 CC DR HSSP; P80378; IAB3.  
 CC DR InterPro; IPR000589; Ribosomal\_S15.  
 CC DR InterPro; IPR005290; RS15\_bact.  
 CC DR Pfam; PF00312; Ribosomal\_S15; 1.  
 CC DR ProDom; PD157043; RS15\_bact; 1.  
 CC DR TIGRFAMs; TIGR00952; S15\_bact; 1.  
 CC DR PROSITE; PS00362; RIBOSOMAL\_S15; 1.  
 CC KW Ribosomal protein; Chloroplast; rRNA-binding.  
 CC SEQUENCE 90 AA; 10904 MW; EC6AA932D6D7EESA CRC64;  
 CC -----  
 CC Query Match 31.7%; Score 32; DB 1; Length 90;  
 CC Best Local Similarity 75.0%; Pred. No. 2.9e+02;  
 CC Matches 6; Conservative 2; Mismatches 0; Indels 0; Gaps 0;  
 CC -----  
 CC QY 8 IISKEKE 15  
 CC Db :||:||||  
 CC 9 VISQEKKE 16  
 CC -----  
 CC RESULT 32  
 CC ID\_R15\_SPIOL STANDARD; PRT; 90 AA.  
 CC AC QW314; P82138;  
 CC DT 16-OCT-2001 (Rel. 40, Created)  
 CC DT 16-OCT-2001 (Rel. 40, Last sequence update)  
 CC DT 16-OCT-2001 (Rel. 40, Last annotation update)  
 CC DE Chloroplast 30S ribosomal protein S15.  
 CC GN RPS15.  
 CC OS Spinacia oleracea (Spinach).  
 CC OG Chloroplast.  
 CC OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 CC OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
 CC OC Caryophyllidae; Caryophyllales; Chenopodiaceae; Spinacia.  
 CC OX NCBI\_TaxID=3562;  
 CC RN [1]  
 CC RP SEQUENCE FROM N.A.  
 CC RC STRAIN=cv. Geant d'hiver, and cv. Monatol;  
 CC RX MEDLINE=21187424; PubMed=11292076;  
 CC RA Schmitz-Linneweber C., Maier R.M., Alcaraz J.-P., Cottet A.,  
 CC RA Hermann R.G., Mache R.;  
 CC RT "The plastid chromosome of spinach (Spinacia oleracea): complete  
 CC RT nucleotide sequence and gene organization.";  
 CC RL Plant Mol. Biol. 45:307-315(2001).  
 CC RN [2]  
 CC RP SEQUENCE OF 1-20, FUNCTION, AND MASS SPECTROMETRY.  
 CC RC STRAIN=cv. Alvaro; TISSUE=Leaf;  
 CC RX MEDLINE=20435797; PubMed=10874039;  
 CC RA Yamauchi K., von Knoblauch K., Subramanian A.R.;  
 CC RT "The plastid ribosomal proteins. Identification of all the proteins in  
 CC RT the small subunit of an organelle ribosome (chloroplast).";  
 CC RL J. Biol. Chem. 37:28455-28465(2000).  
 CC CC -1- FUNCTION: Binds directly to 16S ribosomal RNA.  
 CC CC -1- SUBCELLULAR LOCATION: Chloroplast.  
 CC CC -1- TISSUE SPECIFICITY: Expressed in all plant tissues.  
 CC CC -1- MASS SPECTROMETRY: MW=10830.1; METHOD=Electrospray.  
 CC CC -1- MASS SPECTROMETRY: MW=10847; METHOD=MALDI.  
 CC CC -1- SIMILARITY: BELONGS TO THE S15P FAMILY OF RIBOSOMAL PROTEINS.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial

CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 CC EMBL; AJ400848; CAB88791.1; --  
 CC DR HSSP; P80378; IAB3.  
 CC DR InterPro; IPR000589; Ribosomal\_S15.  
 CC DR InterPro; IPR005290; RS15\_bact.  
 CC DR Pfam; PF00312; Ribosomal\_S15; 1.  
 CC DR ProDom; PD157043; RS15\_bact; 1.  
 CC DR TIGRFAMs; TIGR00952; S15\_bact; 1.  
 CC DR PROSITE; PS00362; RIBOSOMAL\_S15; 1.  
 CC KW Ribosomal protein; Chloroplast; rRNA-binding.  
 CC FT CONFLICT 12 D -> P (IN REF. 2).  
 CC SEQUENCE 90 AA; 10754 MW; F93B9E70B1310413 CRC64;  
 CC -----  
 CC Query Match 31.7%; Score 32; DB 1; Length 90;  
 CC Best Local Similarity 60.0%; Pred. No. 2.9e+02;  
 CC Matches 6; Conservative 3; Mismatches 1; Indels 0; Gaps 0;  
 CC -----  
 CC QY 6 MNIISKEKE 15  
 CC Db :||:||||  
 CC 7 ISVISDEKKE 16  
 CC -----  
 CC RESULT 33  
 CC ID\_Y530\_BUCAI STANDARD; PRT; 90 AA.  
 CC AC P57596;  
 CC DT 16-OCT-2001 (Rel. 40, Created)  
 CC DT 16-OCT-2001 (Rel. 40, Last sequence update)  
 CC DT 16-OCT-2001 (Rel. 40, Last annotation update)  
 CC DE Hypothetical protein BUS30.  
 CC GN BU530.  
 CC OS Buchnera aphidicola (subsp. Acyrthosiphon pisum) (Acyrthosiphon pisum  
 CC symbiotic bacterium).  
 CC OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;  
 CC OC Enterobacteriaceae; Buchnera.  
 CC OX NCBI\_TaxID=118099;  
 CC RN [1]  
 CC RP SEQUENCE FROM N.A.  
 CC RC STRAIN=Tokyo 1998;  
 CC RX MEDLINE=20445173; PubMed=10993077;  
 CC RA Shigenobu S., Watanabe H., Hattori M., Sakaki Y., Ishikawa H.;  
 CC RT "Genome sequence of the endocellular bacterial symbiont of aphids  
 CC RT Buchnera sp. APS.";  
 CC RL Nature 407:81-86(2000).  
 CC CC -1- SIMILARITY: STRONG, TO E. COLI YHEE.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 CC EMBL; AF001119; BAB13223.1; --  
 CC DR Pfam; PF04077; DsrH; 1.  
 CC KW Hypothetical protein; Complete proteome.  
 CC SEQUENCE 90 AA; 10501 MW; 4D9FF2FDB25E338F CRC64;  
 CC -----  
 CC Query Match 31.7%; Score 32; DB 1; Length 90;  
 CC Best Local Similarity 27.8%; Pred. No. 2.9e+02;  
 CC Matches 5; Conservative 6; Mismatches 7; Indels 0; Gaps 0;  
 CC -----  
 CC QY 1 NVLMANNIISKEKEIKW 18  
 CC Db | : : : | : : : |  
 CC 73 NYIHFVSLTKHKQMTW 90  
 CC -----  
 CC RESULT 34  
 CC ID\_YHBY\_ECOLI STANDARD; PRT; 97 AA.



AC P42550;  
 DT 01-NOV-1995 (Rel. 32, Last sequence update)  
 DT 01-NOV-1995 (Rel. 32, Last sequence update)  
 DT 28-FEB-2003 (Rel. 41, Last annotation update)  
 DE Hypothetical protein ybbY.  
 GN YBBY OR B3180 OR C3937 OR Z4542 OR ECS4059 OR SF3220.  
 OS Escherichia coli.  
 OS Escherichia coli O6.  
 OS Escherichia coli O157:H7, and  
 OS Shigella flexneri.  
 OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;  
 OC Enterobacteriaceae; Escherichia  
 OX NCBI\_TaxID=562, 217992, 83334, 623;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC SPECIES=E.coli; STRAIN=K12 / MG1655;  
 RX MEDLINE=97426617; PubMed=9278503;  
 RA Blattner F.R., Plunkett G. III, Bloch C.A., Perna N.T., Burland V.,  
 RA Riley M., Collado-Vides J., Glasner J.D., Rode C.K., Mayhew G.F.,  
 RA Gregor J., Davis N.W., Kirkpatrick H.A., Goeden M.A., Rose D.J.,  
 RA Mau B., Shao Y.;  
 RT "The complete genome sequence of Escherichia coli K-12.";  
 RL Science 277:1453-1474 (1997).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC SPECIES=E.coli; STRAIN=K12;  
 RA Wang R., Kushner S.R.;  
 RL Submitted (SEP-1993) to the EMBL/GenBank/DBJ databases.  
 RN [3]  
 RP SEQUENCE FROM N.A.  
 RC SPECIES=E.coli; STRAIN=O6:H1 / CFT073 / ATCC 700928;  
 RX MEDLINE=22388234; PubMed=12471157;  
 RA Welch R.A., Burland V., Plunkett G. III, Redford P., Roesch P.,  
 RA Rasko D., Buckles E.L., Liou S.-R., Boutin A., Hackett J., Stroud D.,  
 RA Mayhew G.F., Rose D.J., Zhou S., Schwartz D.C., Perna N.T.,  
 RA Mobley H.L.T., Domeneberg M.S., Blattner F.R.;  
 RT "Extensive mosaic structure revealed by the complete genome sequence  
 of uropathogenic Escherichia coli.";  
 RL Proc. Natl. Acad. Sci. U.S.A. 99:17020-17024 (2002).  
 RN [4]  
 RP SEQUENCE FROM N.A.  
 RC SPECIES=E.coli; STRAIN=O157:H7 / EDL933 / ATCC 700927;  
 RX MEDLINE=21074935; PubMed=11206551;  
 RA Perna N.T., Plunkett G. III, Burland V., Mau B., Glasner J.D.,  
 RA Rose D.J., Mayhew G.F., Evans P.S., Gregor J., Kirkpatrick H.A.,  
 RA Posfai G., Hackett J., Klink S., Boutin A., Shao Y., Miller L.,  
 RA Grobeck E.J., Davis N.W., Lim A., Dimalanta E.T., Potamouais K.,  
 RA Apodaca J., Anantharaman T.S., Lin J., Yen G., Schwartz D.C.,  
 RA Welch R.A., Blattner F.R.;  
 RT "Genome sequence of enterohaemorrhagic Escherichia coli O157:H7.";  
 RL Nature 409:529-533 (2001).  
 RN [5]  
 RP SEQUENCE FROM N.A.  
 RC SPECIES=E.coli; STRAIN=O157:H7 / RIMD 0509952;  
 RX MEDLINE=21156231; PubMed=11258796;  
 RA Hayashi T., Makino K., Ohnishi M., Kurokawa K., Ishii K., Yokoyama K.,  
 RA Han C.-G., Ohtsubo E., Nakayama K., Murata T., Tanaka M., Tobe T.,  
 RA Tada T., Takami H., Honda T., Sasakawa C., Ogasawara N., Yasunaga T.,  
 RA Kuhara S., Shiba T., Hattori M., Shinegawa H.;  
 RT "Complete genome sequence of enterohemorrhagic Escherichia coli  
 O157:H7 and genomic comparison with a laboratory strain K-12.";  
 RL Nature Res. 8:11-22 (2001).  
 RN [6]  
 RP SEQUENCE OF 64-97 FROM N.A.  
 RC SPECIES=E.coli;  
 RX MEDLINE=91057148; PubMed=2243801;  
 RA Sparkowski J., Das A.;  
 RT "The nucleotide sequence of greA, a suppressor gene that restores  
 growth of an Escherichia coli RNA polymerase mutant at high  
 temperature.";  
 RL Nucleic Acids Res. 18:6443-6443 (1990).  
 RN [7]  
 RP SEQUENCE FROM N.A.

RC SPECIES=S.flexneri; STRAIN=301 / Serotype 2a;  
 RX MEDLINE=22272406; PubMed=12384590;  
 RA Jin Q., Yuan Z., Xu J., Wang Y., Shen Y., Lu W., Wang J., Liu H.,  
 RA Yang J., Yang F., Zhang X., Zhang J., Yang G., Wu H., Qu D., Dong J.,  
 RA Sun L., Xue Y., Zhao A., Gao Y., Zhu J., Kan B., Ding K., Chen S.,  
 RA Cheng H., Yao Z., He B., Chen R., Ma D., Qiang B., Wen Y., Hou Y.,  
 RA Yu J.;  
 RT "Genome sequence of Shigella flexneri 2a: insights into pathogenicity  
 through comparison with genomes of Escherichia coli K12 and O157.";  
 RL Nucleic Acids Res. 30:4432-4441 (2002).  
 CC -1- SIMILARITY: BELONGS TO THE UPF0044 FAMILY. STRONG, TO H.INFLUENZA  
 CC HI1333.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 CC EMBL; UI18997; AAA57981.1; -;  
 DR EMBL; AE000398; AAC76212.1; -;  
 DR EMBL; U01376; -, NOT\_ANNOTATED\_CDS.  
 DR EMBL; AE016767; AAN82377.1; -;  
 DR EMBL; AE005546; AAG58314.1; -;  
 DR EMBL; AP002564; BAB37482.1; -;  
 DR EMBL; X54718; -, NOT\_ANNOTATED\_CDS.  
 DR EMBL; AE015333; AAN44686.1; -;  
 DR PIR; C91136; C91136.  
 DR PIR; F65108; F65108.  
 DR PIR; F85981; F85981.  
 DR EcoGene; EG12794; YbbY.  
 DR InterPro; IPR001890; UPF0044.  
 DR Pfam; PF01985; UPF0044; 1.  
 DR ProDom; PD010559; UPF0044; 1.  
 DR TIGRFAMs; TIGR00253; TIGR00253; 1.  
 DR PROSITE; PS01301; UPF0044; 1.  
 KW Hypothetical protein; Complete proteome.  
 SQ SEQUENCE 97 AA; 10784 MW; 52B4F7AD4C203382 CRC64;  
 Query Match 31.7%; Score 32; DB 1; Length 97;  
 Best Local Similarity 41.7%; Pred. No. 3.1e+02;  
 Matches 5; Conservative 5; Mismatches 2; Indels 0; Gaps 0;  
 QY 6 MNTISKEKEIK 17  
 DB 1 MNLSTKQKHLK 12  
 ||: :||: :||  
 RESULT 35  
 Y126 METJA STANDARD; PRT; 98 AA.  
 AC Q57590;  
 DT 01-NOV-1997 (Rel. 35, Created)  
 DT 15-DEC-1998 (Rel. 37, Last sequence update)  
 DT 16-OCT-2001 (Rel. 40, Last annotation update)  
 DE Hypothetical protein MJ0126.  
 GN MJ0126.  
 OS Methanococcus jannaschii.  
 OC Archaea; Euryarchaeota; Methanococci; Methanococcales;  
 OC Methanocaldococcaceae; Methanocaldococcus.  
 OX NCBI\_TaxID=2190;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=JAL-1 / DSM 2661 / ATCC 43067;  
 RX MEDLINE=96337999; PubMed=8688087;  
 RA Bult C.J., White O., Olsen G.J., Zhou L., Fleischmann R.D.,  
 RA Sutton G.G., Blake J.A., FitzGerald L.M., Clayton R.A., Gocayne J.D.,  
 RA Kerlavage A.R., Dougherty B.A., Tomb J.-F., Adams M.D., Reich C.I.,  
 RA Overbeek R., Kirkness E.F., Weinstock K.G., Merrick J.M., Glodek A.,  
 RA Scott J.L., Geoghagen N.S.M., Weidman J.F., Fuhrmann J.L., Nguyen D.,  
 RA Utterback T.R., Kelley J.M., Peterson J.D., Sadow P.W., Hanna M.C.,

```
RA Cotton M.D., Roberts K.M., Hurst M.A., Kaine B.P., Borodovsky M.,
RA Klenk H.-P., Fraser C.M., Smith H.O., Woese C.R., Venter J.C.;
RT "Complete genome sequence of the methanogenic archaeon, Methanococcus
RT jannaschii."
RL Science 273:1058-1073(1996).
CC -1- SIMILARITY: BELONGS TO THE M.JANNASCHII MJ0126 / MJ0128 / MJ0141 /
CC MJ0435 / MJ0604 / MJ1215 / MJ1305 / MJ1379 FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; U67469; AAB98106.1; -.
CC TIGR; MJ0126; -.
CC InterPro; IPR002934; NTP transf.
CC Pfam; PF01909; NTP transf_2; 1.
CC Hypothetical protein; Complete proteome.
CC SEQUENCE 98 AA; 11604 MW; 0748F07131F29695 CRC64;
CC -----
CC Query Match 31.7%; Score 32; DB 1; Length 98;
CC Best Local Similarity 60.0%; Pred. No. 3.1e+02;
CC Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;
CC -----
QY 8 IISKEKKEIK 17
Db 10 ILRKHKKELK 19
| : |||||
| : |||||
-----
RESULT 36
RR14 ODOSI STANDARD; PRT; 100 AA.
AC P49502;
DT 01-FEB-1996 (Rel. 33, Created)
DT 01-FEB-1996 (Rel. 33, Last sequence update)
DT 01-FEB-1996 (Rel. 33, Last annotation update)
DE Chloroplast 30S ribosomal protein S14.
GN RPS14.
OS Odontella sinensis (Marine centric diatom).
OG Chloroplast.
OC Eukaryota; stramenopiles; Bacillariophyta; Coscinodiscophyceae;
OC Biddulphiophycidae; Eupodiscales; Eupodisaceae; Odontella.
OX NCBI_TaxID=2839;
RN [1]
RP SEQUENCE FROM N.A.
RA Kowalik K.V., Stoebe B., Schaffran I., Kroth-Pancic P., Freier U.;
RT "The chloroplast genome of a chlorophyll a+c-containing alga,
RT Odontella sinensis."
RL Plant Mol. Biol. Rep. 13:336-342(1995).
CC -1- SIMILARITY: BELONGS TO THE S14P FAMILY OF RIBOSOMAL PROTEINS.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; Z67753; CAA91675.1; -.
CC PIR; S78302; S78302.
CC InterPro; IPR001209; Ribosomal S14.
CC Pfam; PF00253; Ribosomal S14; 1.
CC PROSITE; PS00527; RIBOSOMAL_S14; 1.
CC Ribosomal protein; Chloroplast.
CC SEQUENCE 100 AA; 11988 MW; E21AB650D567A74E CRC64;
CC -----
CC Query Match 31.7%; Score 32; DB 1; Length 100;
CC Best Local Similarity 54.5%; Pred. No. 3.2e+02;
CC Matches 6; Conservative 3; Mismatches 2; Indels 0; Gaps 0;
```

```
QY 7 NIISKEKKEIK 17
Db 5 SMIEREKKRIK 15
| : |||||
| : |||||
-----
RESULT 37
YF50 METJA STANDARD; PRT; 61 AA.
AC Q58945;
DT 15-JUL-1998 (Rel. 36, Created)
DT 15-JUL-1998 (Rel. 36, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Hypothetical protein MJ1550.
GN MJ1550.
OS Methanococcus jannaschii.
OC Archaea; Euryarchaeota; Methanococci; Methanococcales;
OC Methanocaldococcales; Methanocaldococcus.
OX NCBI_TaxID=2190;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=JAL-1 / DSM 2661 / ATCC 43067;
RX MEDLINE=96337999; PubMed=8688087;
RA Bult C.J., White O., Olsen G.J., Zhou L., Fleischmann R.D.,
RA Sutton G.G., Blake J.A., Fitzgerald L.M., Clayton R.A., Gocayne J.D.,
RA Kerlavage A.R., Dougherty B.A., Tomb J.-P., Adams M.D., Reich C.I.,
RA Overbeek R., Kirkness E.F., Weinstock K.G., Merrick J.M., Glodek A.,
RA Scott J.L., Geoghegan N.S.M., Weidman J.P., Fuhrmann J.L., Nguyen D.,
RA Utterback T.R., Kelley J.M., Peterson J.D., Sadow P.W., Hanna M.C.,
RA Cotton M.D., Roberts K.M., Hurst M.A., Kaine B.P., Borodovsky M.,
RA Klenk H.-P., Fraser C.M., Smith H.O., Woese C.R., Venter J.C.;
RT "Complete genome sequence of the methanogenic archaeon, Methanococcus
RT jannaschii."
RL Science 273:1058-1073(1996).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; U67596; AAB99580.1; -.
CC PIR; E64493; E64493.
CC TIGR; MJ1550; -.
CC Hypothetical protein; Complete proteome.
CC SEQUENCE 61 AA; 7106 MW; 52C0B8011E728F12 CRC64;
CC -----
CC Query Match 30.7%; Score 31; DB 1; Length 61;
CC Best Local Similarity 41.7%; Pred. No. 2.8e+02;
CC Matches 5; Conservative 3; Mismatches 4; Indels 0; Gaps 0;
CC -----
QY 9 ISKEKKEIKWIG 20
Db 11 ICKKCKCVMWVG 22
| : |||||
| : |||||
-----
RESULT 38
GBGU BOVIN STANDARD; PRT; 69 AA.
ID GBGU BOVIN
AC P50154;
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Guanine nucleotide-binding protein G(1)/G(S)/G(O) gamma-T2 subunit
DE (G gamma-C) (G-gamma-8).
GN GNGT2 OR GNG8 OR GNGT8.
OS Bos taurus (Bovine).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
OC Bovidae; Bovinae; Bos.
OX NCBI_TaxID=9913;
```

RN SEQUENCE FROM N.A.  
 RP TISSUE=Retina;  
 RA MEDLINE=92238332; PubMed=7721746;  
 RX Ong O.C., Yamane H.K., Phan K.B., Fong H.K., Bok D., Lee R.H.,  
 RA Fung B.K.-K.;  
 RT "Molecular cloning and characterization of the G protein gamma  
 subunit of cone photoreceptors.";  
 RL J. Biol. Chem. 270:8495-8500(1995).  
 CC -1- FUNCTION: GUANINE NUCLEOTIDE-BINDING PROTEINS (G PROTEINS) ARE  
 INVOLVED AS A MODULATOR OR TRANSDUCER IN VARIOUS TRANSMEMBRANE  
 SIGNALING SYSTEMS. THE BETA AND GAMMA CHAINS ARE REQUIRED FOR THE  
 GTPASE ACTIVITY, FOR REPLACEMENT OF GDP BY GTP, AND FOR G PROTEIN-  
 EFFECTOR INTERACTION.  
 CC -1- SUBUNIT: G proteins are composed of 3 units (alpha, beta and  
 gamma).  
 CC -1- SIMILARITY: BELONGS TO THE G PROTEIN GAMMA FAMILY.  
 CC  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 the European Bioinformatics Institute. There are no restrictions on its  
 use by non-profit institutions as long as its content is in no way  
 modified and this statement is not removed. Usage by and for commercial  
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC  
 CC EMBL; U20085; AAB61306.1; -.  
 DR PIR; A56378; A56378.  
 DR HSP; P02698; IAOE.  
 DR InterPro; IPR001770; G-gamma.  
 DR Pfam; PF00631; G-gamma; 1.  
 DR PRINTS; PR00321; GPROTEIN.  
 DR ProDom; PD003783; G-gamma; 1.  
 DR SMART; SM00224; GGU; 1.  
 DR PROSITE; PS50058; G PROTEIN GAMMA; 1.  
 KW Transducer; Prenylation; Lipoprotein; Multigene family.  
 FT LIPID 66 66 FARNESYL (BY SIMILARITY).  
 FT PROPEP 67 69 REMOVED IN MATURE FORM (BY SIMILARITY).  
 SQ SEQUENCE 69 AA; 7728 MW; 210C8319D1520314 CRC64;  
  
 Query Match 30.7%; Score 31; DB 1; Length 69;  
 Best Local Similarity 70.0%; Pred. No. 3.2e+02;  
 Matches 7; Conservative 1; Mismatches 2; Indels 0; Gaps 0;  
  
 QY 8 IISKEKKEIK 17  
 :|||:|  
 DB 28 LISKTGKEIK 37  
  
 RESULT 39  
 ACPH\_UREPA STANDARD; PRT; 77 AA.  
 ID\_ACPH\_UREPA  
 AC QPPYA;  
 DT 28-FEB-2003 (Rel. 41, Created)  
 DT 28-FEB-2003 (Rel. 41, Last sequence update)  
 DT 15-SEP-2003 (Rel. 42, Last annotation update)  
 DE Acyl carrier protein homolog (ACP).  
 GN U0506.  
 OS Ureaplasma parvum (Ureaplasma urealyticum biotype 1).  
 OC Bacteria; Firmicutes; Mollicutes; Mycoplasmataceae; Ureaplasma.  
 OX NCBI\_TaxID=134821;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=Serovar 3;  
 RX MEDLINE=20500219; PubMed=11048724;  
 RA Glass J.I., Lefkowitz E.J., Glass J.S., Heiner C.R., Chen E.Y.,  
 RA Cassell G.H.;  
 RT "The complete sequence of the mucosal pathogen Ureaplasma  
 urealyticum.";  
 RL Nature 407:757-762(2000).  
 CC -1- FUNCTION: Carrier of the growing fatty acid chain in fatty acid  
 biosynthesis (By similarity).  
 CC -1- PATHWAY: De novo fatty acid biosynthesis.

CC -1- PTM: 4'-phosphopantetheine is transferred from CoA to a specific  
 serine of the apo-ACP-like protein (Potential).  
 CC -1- SIMILARITY: Contains 1 acyl carrier domain.  
 CC  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 the European Bioinformatics Institute. There are no restrictions on its  
 use by non-profit institutions as long as its content is in no way  
 modified and this statement is not removed. Usage by and for commercial  
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC  
 CC EMBL; AE002149; AAF30918.1; -.  
 DR InterPro; IPR006163; Pp bind.  
 DR InterPro; IPR006162; Ppantenn attach.  
 DR Pfam; PF00550; pp-binding; 1.  
 DR PROSITE; PS50075; ACP\_DOMAIN; 1.  
 DR PROSITE; PS00012; PHOSPHOPANTETHEINE; 1.  
 KW Hypothetical protein; Fatty acid biosynthesis; Phosphopantetheine;  
 Complete proteome.  
 FT BINDING 36 36 PHOSPHOPANTETHEINE (BY SIMILARITY).  
 SQ SEQUENCE 77 AA; 8750 MW; 3CDB655FBFA968C6 CRC64;  
  
 Query Match 30.7%; Score 31; DB 1; Length 77;  
 Best Local Similarity 50.0%; Pred. No. 3.5e+02;  
 Matches 7; Conservative 3; Mismatches 4; Indels 0; Gaps 0;  
  
 QY 3 LMAWNLISKEKEIK 16  
 :|||:|  
 DB 37 LSAWNLIMKIEDQI 50  
  
 RESULT 40  
 THIO\_METTM STANDARD; PRT; 83 AA.  
 ID\_THIO\_METTM  
 AC P42035;  
 DT 01-NOV-1995 (Rel. 32, Created)  
 DT 01-NOV-1995 (Rel. 32, Last sequence update)  
 DT 16-OCT-2001 (Rel. 40, Last annotation update)  
 DE Probable Thioredoxin (Glutaredoxin-like protein).  
 OS Methanobacterium thermoautotrophicum (strain Marburg / DSM 2133).  
 OC Archaea; Euryarchaeota; Methanobacteria; Methanobacteriales;  
 OC Methanobacteriaceae; Methanothermobacter.  
 OX NCBI\_TaxID=79929;  
 RN [1]  
 RP SEQUENCE  
 RX MEDLINE=92268099; PubMed=1587836;  
 RA McFarlan S.C., Terrell C.A., Hogenkamp H.P.C.;  
 RT "The purification, characterization, and primary structure of a small  
 redox protein from Methanobacterium thermoautotrophicum, an  
 archaeobacterium.";  
 RL J. Biol. Chem. 267:10561-10569(1992).  
 CC -1- FUNCTION: DOES NOT FUNCTION AS A GLUTATHIONE-DISULFIDE  
 OXIDOREDUCTASE IN THE PRESENCE OF GLUTATHIONE AND GLUTATHIONE  
 REDUCTASE. MAY BE A COMPONENT OF A RIBONUCLEOTIDE-REDUCING SYSTEM  
 DISTINCT FROM THE PREVIOUSLY DESCRIBED SYSTEMS UTILIZING  
 THIOREDOXIN OR GLUTAREDOXIN.  
 CC -1- SUBCELLULAR LOCATION: Cytoplasmic.  
 CC -1- SIMILARITY: BELONGS TO THE GLUTAREDOXIN FAMILY.  
 DR InterPro; IPR002109; Glutaredoxin.  
 DR InterPro; IPR004502; Thio glut.  
 DR Pfam; PF006663; Thioredox dom2.  
 DR Pfam; PF00462; Glutaredoxin; 1.  
 DR TIGRFAMs; TIGR00411; redox disulf. 1; 1.  
 DR PROSITE; PS00195; GLUTAREDOXIN; 1.  
 KW Redox-active center; Electron transport.  
 FT DISULFID 12 15 REDOX-ACTIVE (BY SIMILARITY).  
 SQ SEQUENCE 83 AA; 9147 MW; F538E8023A83F800 CRC64;  
  
 Query Match 30.7%; Score 31; DB 1; Length 83;  
 Best Local Similarity 41.7%; Pred. No. 3.8e+02;  
 Matches 5; Conservative 4; Mismatches 3; Indels 0; Gaps 0;

QY 4 MAMNIISKEKE 15  
 ||: : : |||  
 DB 17 MAIEVVDEAKE 28

## RESULT 41

THIO\_METH STANDARD; PRT; 84 AA.

AC O26898;  
 DT 15-JUL-1998 (Rel. 36, Created)  
 DT 16-OCT-2001 (Rel. 40, Last sequence update)  
 DT 16-OCT-2001 (Rel. 40, Last annotation update)  
 DE Probable Thioredoxin (Glutaredoxin-like protein).  
 GN MTH807.  
 OS Methanobacterium thermoautotrophicum.  
 OC Archaea; Euryarchaeota; Methanobacteria; Methanobacteriales;  
 OC Methanobacteriaceae; Methanothermobacter.  
 OX NCBI\_TaxID=187420;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=Delta H;  
 RX MEDLINE=98037514; PubMed=9371463;  
 RA Smith D.R., Doucette-Stamm L.A., Deloughery C., Lee H.-M., Dubois J.,  
 RA Aldredge T., Bashirzadeh R., Blakely D., Cook R., Gilbert K.,  
 RA Harrigan D., Hoang L., Keagle P., Lumm W., Pothier B., Qiu D.,  
 RA Spadafora R., Vicare R., Wang Y., Wierzbowski J., Gibson R.,  
 RA Jiwan N., Caruso A., Bush D., Safer H., Patwell D., Prabhakar S.,  
 RA McDougall S., Shimer G., Goyal A., Pietrowski S., Church G.M.,  
 RA Daniels C.J., Mao J.-I., Rice P., Noelling J., Reeve J.N.;  
 RT "Complete genome sequence of Methanobacterium thermoautotrophicum  
 RT deltaH: functional analysis and comparative genomics.";  
 RL J. Bacteriol. 179:7135-7155(1997).  
 CC -1- FUNCTION: ACTS TO MAINTAIN REDOX HOMEOSTASIS; FUNCTIONS AS A  
 CC PROTEIN DISULFIDE REDUCTASE (BY SIMILARITY).  
 CC -1- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).  
 CC -1- SIMILARITY: BELONGS TO THE GLUTAREDOXIN FAMILY.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL; AE000858; AAB85307.1; --  
 DR PIR; G69207; G69207.  
 DR InterPro; IPR002109; Glutaredoxin.  
 DR InterPro; IPR004502; Thio glut.  
 DR InterPro; IPR006663; Thioredox dom2.  
 DR Pfam; PF00462; glutaredoxin; 1-  
 DR TIGRFAMs; TIGR00411; redox\_disulf\_1; 1.  
 DR PROSITE; PS00195; GLUTAREDOXIN; 1.  
 DR Redox-active center; Electron transport; Complete proteome.  
 KW INIT MET 0 BY SIMILARITY.  
 FT DISULFID 12 15 REDOX-ACTIVE (BY SIMILARITY).  
 FT SEQUENCE 84 AA; 9350 MW; 3B2F48EF0E717145 CRC64;  
 SQ

Query Match 30.7%; Score 31; DB 1; Length 84;  
 Best Local Similarity 41.7%; Pred. No. 3.9e+02;  
 Matches 5; Conservative 4; Mismatches 3; Indels 0; Gaps 0;

QY 4 MAMNIISKEKE 15  
 ||: : : |||  
 DB 17 MAIEVVDEAKE 28

## RESULT 42

ID\_CYC6 SYNLI STANDARD; PRT; 87 AA.

AC P00114;  
 DT 21-JUL-1986 (Rel. 01, Created)  
 DT 21-JUL-1986 (Rel. 01, Last sequence update)

DT 28-FEB-2003 (Rel. 41, Last annotation update)  
 DE Cytochrome c6 (Soluble cytochrome f) (Cytochrome c553) (Cytochrome c-  
 DE 553).  
 GN PETJ.  
 OS Synechococcus lividus.  
 OC Bacteria; Cyanobacteria; Chroococcales; Synechococcus.  
 OX NCBI\_TaxID=33070;  
 RN [1]  
 RP SEQUENCE.  
 RA Borden D., Margolias E.;  
 RL Submitted (DSC-1979) to the PIR data bank.  
 CC -1- FUNCTION: Functions as an electron carrier between membrane-bound  
 CC cytochrome b6f and photosystem I in oxygenic photosynthesis (By  
 CC similarity).  
 CC -1- SUBUNIT: Monomer (By similarity).  
 CC -1- SUBCELLULAR LOCATION: Cytoplasmic (Potential).  
 CC -1- PTM: Binds one heme group per molecule.  
 CC -1- SIMILARITY: Belongs to the cytochrome c family. PetJ subfamily.  
 DR PIR; A00106; CCYC6L.  
 DR HSP; P56534; IC6S.  
 DR HAMAP; MF\_00594; -; 1.  
 DR InterPro; IPR003088; Cyt CI.  
 DR InterPro; IPR002329; Cyt\_CIC.  
 DR InterPro; IPR000345; CytC\_heme\_bind.  
 DR Pfam; PF00034; cytochrome\_c; 1-  
 DR PRINTS; PR00605; CYCHROME\_C; 1.  
 DR PROSITE; PS00190; CYTOCHROME\_C; 1.  
 KW Electron transport; Photosynthesis; Heme.  
 FT BINDING 14 14 HEME (COVALENT).  
 FT BINDING 17 17 HEME (COVALENT).  
 FT BINDING 18 18 IRON (HEME AXIAL LIGAND).  
 FT METAL 58 58 IRON (HEME AXIAL LIGAND).  
 FT SEQUENCE 87 AA; 9129 MW; 37713EC6405EBEE CRC64;  
 SQ

Query Match 30.7%; Score 31; DB 1; Length 87;  
 Best Local Similarity 46.7%; Pred. No. 4e+02;  
 Matches 7; Conservative 2; Mismatches 6; Indels 0; Gaps 0;

QY 1 NVLMAMNIISKEKE 15  
 ||: : : |||  
 DB 23 NVVMANKTLKKEALE 37

## RESULT 43

CYC6\_SYNPN3 STANDARD; PRT; 87 AA.

AC P00115;  
 DT 21-JUL-1986 (Rel. 01, Created)  
 DT 21-JUL-1986 (Rel. 01, Last sequence update)  
 DT 28-FEB-2003 (Rel. 41, Last annotation update)  
 DE Cytochrome c6 (Soluble cytochrome f) (Cytochrome c553) (Cytochrome c-  
 DE 553) (Cytochrome c-552).  
 GN PETJ.

OS Synechococcus sp. (strain PCC 6312 / ATCC 27167).  
 OC Bacteria; Cyanobacteria; Chroococcales; Synechococcus.  
 OX NCBI\_TaxID=195253;  
 RN [1]  
 RP SEQUENCE.

RX MEDLINE=80068924; PubMed=228936;

RA Aitken A.;  
 RL "Purification and primary structure of cytochrome c-552 from the  
 RT cyanobacterium, Synechococcus PCC 6312.";  
 RL Eur. J. Biochem. 101:297-308(1979).

CC -1- FUNCTION: Functions as an electron carrier between membrane-bound  
 CC cytochrome b6f and photosystem I in oxygenic photosynthesis.  
 CC -1- SUBUNIT: Monomer (By similarity).  
 CC -1- SUBCELLULAR LOCATION: Cytoplasmic (Potential).

CC -1- PTM: Binds one heme group per molecule.  
 CC -1- SIMILARITY: Belongs to the cytochrome c family. PetJ subfamily.

DR HSP; P56534; IC6S.

DR HAMAP; MF\_00594; -; 1.

DR InterPro; IPR003088; Cyt CI.

DR InterPro; IPR002329; Cyt\_CIC.

```
DR InterPro; IPR000345; CytC_heme_bind.
DR Pfam; PF00034; cytochrome c; 1.
DR PRINTS; P00605; CYTOCHROME C.
DR PROSITE; PS00190; CYTOCHROME C; 1.
KW Electron transport; Photosynthesis; Heme.
FT BINDING 14 14 HEME (COVALENT).
FT BINDING 17 17 HEME (COVALENT).
FT METAL 18 18 IRON (HEME AXIAL LIGAND).
FT METAL 58 58 IRON (HEME AXIAL LIGAND).
SQ SEQUENCE 87 AA; 9098 MW; 41FC9734E3DF83F2 CRC64;

Query Match 30.7%; Score 31; DB 1; Length 87;
Best Local Similarity 46.7%; Pred. No. 4e+02;
Matches 7; Conservative 2; Mismatches 6; Indels 0; Gaps 0;

QY 1 NVLMAMNIISKEKKE 15
DB 23 NVVMANKTLKKEALE 37

RESULT 44
CY2_RHOTE STANDARD; PRT; 92 AA.
AC P00098;
DT 21-JUL-1986 (Rel. 01, Created)
DT 21-JUL-1986 (Rel. 01, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Cytochrome C2.
OS Rhodocyclus tenuis (Rhodospirillum tenue).
OC Bacteria; Proteobacteria; Betaproteobacteria; Rhodocyclales;
OC Rhodocyclaceae; Rhodocyclus.
OX NCBI_TaxID=1066;
RN [1]
RP SEQUENCE.
RC STRAIN=DSM 3761;
RX MEDLINE=79199668; PubMed=221823;
RA Ambler R.P., Meyer T.E., Kamen M.D.;
RT "Anomalies in amino acid sequences of small cytochromes c and
cytochromes c' from two species of purple photosynthetic bacteria.";
RL Nature 278:661-662(1979).
CC -!- FUNCTION: CYTOCHROME C2 IS FOUND MAINLY IN PURPLE, NONSULFUR,
PHOTOSYNTHETIC BACTERIA WHERE IT FUNCTIONS AS THE ELECTRON DONOR
TO THE OXIDIZED BACTERIOCHLOROPHYLL IN THE PHOTOPHOSPHORYLATION
PATHWAY. HOWEVER, IT MAY ALSO HAVE A ROLE IN THE RESPIRATORY CHAIN
AND IS FOUND IN SOME NONPHOTOSYNTHETIC BACTERIA.
CC -!- SIMILARITY: THIS SEQUENCE IS MORE CLOSELY RELATED TO THE SEQUENCES
OF CYTOCHROME C551 FROM PSEUDOMONAS AND AZOTOBACTER THAN TO THE
SEQUENCES OF CYTOCHROME C2 FROM OTHER SPECIES OF RHODOSPIRILLUM.
DR PIR; A00090; CCOQ2T.
DR HSSP; P00099; 451C.
DR InterPro; IPR003088; Cyt_C1.
DR Pfam; PF00034; cytochrome_c; 1.
DR PROSITE; PS00190; CYTOCHROME C; 1.
KW Electron transport; Photosynthesis; Heme.
FT BINDING 12 12 HEME (COVALENT).
FT BINDING 15 15 HEME (COVALENT).
FT METAL 16 16 IRON (HEME AXIAL LIGAND).
FT METAL 66 66 IRON (HEME AXIAL LIGAND).
SQ SEQUENCE 92 AA; 9781 MW; 5A51C1E372162F9F CRC64;

Query Match 30.7%; Score 31; DB 1; Length 92;
Best Local Similarity 33.3%; Pred. No. 4.2e+02;
Matches 6; Conservative 6; Mismatches 4; Indels 2; Gaps 1;

QY 4 MAMNIISKEK--KEIKWI 19
DB 66 MPANNVTKEEATRLVKWV 83

RESULT 45
RPOL_METAC STANDARD; PRT; 92 AA.
ID NCBI_TaxID=1066;
RN [1]
RP SEQUENCE.
RC STRAIN=DSM 3761;
RX MEDLINE=79199668; PubMed=221823;
RA Ambler R.P., Meyer T.E., Kamen M.D.;
RT "Anomalies in amino acid sequences of small cytochromes c and
cytochromes c' from two species of purple photosynthetic bacteria.";
RL Nature 278:661-662(1979).
CC -!- FUNCTION: CYTOCHROME C2 IS FOUND MAINLY IN PURPLE, NONSULFUR,
PHOTOSYNTHETIC BACTERIA WHERE IT FUNCTIONS AS THE ELECTRON DONOR
TO THE OXIDIZED BACTERIOCHLOROPHYLL IN THE PHOTOPHOSPHORYLATION
PATHWAY. HOWEVER, IT MAY ALSO HAVE A ROLE IN THE RESPIRATORY CHAIN
AND IS FOUND IN SOME NONPHOTOSYNTHETIC BACTERIA.
CC -!- SIMILARITY: THIS SEQUENCE IS MORE CLOSELY RELATED TO THE SEQUENCES
OF CYTOCHROME C551 FROM PSEUDOMONAS AND AZOTOBACTER THAN TO THE
SEQUENCES OF CYTOCHROME C2 FROM OTHER SPECIES OF RHODOSPIRILLUM.
DR PIR; A00090; CCOQ2T.
DR HSSP; P00099; 451C.
DR InterPro; IPR003088; Cyt_C1.
DR Pfam; PF00034; cytochrome_c; 1.
DR PROSITE; PS00190; CYTOCHROME C; 1.
KW Electron transport; Photosynthesis; Heme.
FT BINDING 12 12 HEME (COVALENT).
FT BINDING 15 15 HEME (COVALENT).
FT METAL 16 16 IRON (HEME AXIAL LIGAND).
FT METAL 66 66 IRON (HEME AXIAL LIGAND).
SQ SEQUENCE 92 AA; 9781 MW; 5A51C1E372162F9F CRC64;

Query Match 30.7%; Score 31; DB 1; Length 92;
Best Local Similarity 35.7%; Pred. No. 4.2e+02;
Matches 5; Conservative 5; Mismatches 4; Indels 0; Gaps 0;

QY 4 MAMNIISKEKKEIK 17
DB 1 MELNLIKTNNELE 14

Search completed: February 17, 2004, 10:57:03
Job time : 7.33663 secs
```

```
AC Q8TS94;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE DNA-directed RNA polymerase subunit L (EC 2.7.7.6).
GN RPOL OR MA0721.
OS Methanosarcina acetivorans.
OC Archaea; Euryarchaeota; Euryarchaeota orders incertae sedis;
OC Methanosarcinales; Methanosarcinaceae; Methanosarcina.
OX NCBI_TaxID=2214;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CZA / ATCC 35395 / DSM 2834;
RX MEDLINE=21929760; PubMed=11932238;
RA Galagan J.E., Nusbaum C., Roy A., Endrizzi M.G., MacDonald P.,
RA FitzHugh W., Calvo S., Engels R., Smirnov S., Atnoor D., Brown A.,
RA Allen N., Naylor J., Stange-Thomann N., DeAtrallano K., Johnson R.,
RA Linton L., McEwan P., McKernan K., Talamas J., Tirrell A., Ye W.,
RA Zimmer A., Barber R.D., Cann I., Graham D.E., Grahame D.A., Guss A.M.,
RA Hedderich R., Ingram-Smith C., Kuettner H.C., Krzycki J.A.,
RA Leigh J.A., Li W., Liu J., Mukhopadhyay B., Reeve J.N., Smith K.,
RA Springer T.A., Umavam L.A., White O., White R.H., de Macario E.C.,
RA Perry J.G., Jarrell K.F., Jing H., Macario A.J.L., Paulsen I.,
RA Pritchett M., Sowers K.R., Swanson R.V., Zinder S.H., Lander E.,
RA Metcalf W.W., Birren B.;
RT "The genome of Methanosarcina acetivorans reveals extensive metabolic
and physiological diversity.";
RL Genome Res. 12:532-542(2002).
CC -!- FUNCTION: DNA-dependent RNA polymerase catalyzes the transcription
of DNA into RNA using the four ribonucleoside triphosphates as
substrates.
CC -!- CATALYTIC ACTIVITY: N nucleoside triphosphate = N diphosphate +
{RNA} (N).
CC -!- SIMILARITY: BELONGS TO THE ARCHAEABACTERIA RPOL / EUKARYOTIC RPBI1/
RPC19 RNA POLYMERASE SUBUNIT FAMILY.
CC -----
CC This SWISS-PROT entry is copyrighted. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).
CC -----
DR EMBL; A010733; AAM04161.1; -.
DR HAMAP; MF 00261; -.
DR InterPro; IPR001306; RNA_pol_L.
DR Pfam; PF01193; RNA_pol_L; 1.
DR PROSITE; PS01154; RNA_POL_L_13KD; FALSE NEG.
KW Transferase; Transcription; DNA-directed RNA polymerase;
KW Complete proteome.
SQ SEQUENCE 92 AA; 10374 MW; B87AE96EAE9EC783 CRC64;

Query Match 30.7%; Score 31; DB 1; Length 92;
Best Local Similarity 35.7%; Pred. No. 4.2e+02;
Matches 5; Conservative 5; Mismatches 4; Indels 0; Gaps 0;

QY 4 MAMNIISKEKKEIK 17
DB 1 MELNLIKTNNELE 14
```

**THIS PAGE BLANK (USPTO)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: February 17, 2004, 10:50:12 ; Search time 30.4158 Seconds  
(without alignments)  
83.497 Million cell updates/sec

**Title:** US-09-900-147-5

Perfect score: 76

Sequence: 1 RVYDALNVLAMNIIIS 16

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1107863 seqs, 158726573 residues

Total number of hits satisfying chosen parameters: 717921

```
Minimum DB seq length: 0
Maximum DB seq length: 100
```

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 s

Database : A Geneseq 19Jun03:\*

1:	/SIDS1/cgcdata/geneseq/geneseqp-emb1/AA1980.DAT.*
2:	/SIDS1/cgcdata/geneseq/geneseqp-emb1/AA1981.DAT.*
3:	/SIDS1/cgcdata/geneseq/geneseqp-emb1/AA1982.DAT.*
4:	/SIDS1/cgcdata/geneseq/geneseqp-emb1/AA1983.DAT.*
5:	/SIDS1/cgcdata/geneseq/geneseqp-emb1/AA1984.DAT.*
6:	/SIDS1/cgcdata/geneseq/geneseqp-emb1/AA1985.DAT.*
7:	/SIDS1/cgcdata/geneseq/geneseqp-emb1/AA1986.DAT.*
8:	/SIDS1/cgcdata/geneseq/geneseqp-emb1/AA1987.DAT.*
9:	/SIDS1/cgcdata/geneseq/geneseqp-emb1/AA1988.DAT.*
10:	/SIDS1/cgcdata/geneseq/geneseqp-emb1/AA1989.DAT.*
11:	/SIDS1/cgcdata/geneseq/geneseqp-emb1/AA1990.DAT.*
12:	/SIDS1/cgcdata/geneseq/geneseqp-emb1/AA1991.DAT.*
13:	/SIDS1/cgcdata/geneseq/geneseqp-emb1/AA1992.DAT.*
14:	/SIDS1/cgcdata/geneseq/geneseqp-emb1/AA1993.DAT.*
15:	/SIDS1/cgcdata/geneseq/geneseqp-emb1/AA1994.DAT.*
16:	/SIDS1/cgcdata/geneseq/geneseqp-emb1/AA1995.DAT.*
17:	/SIDS1/cgcdata/geneseq/geneseqp-emb1/AA1996.DAT.*
18:	/SIDS1/cgcdata/geneseq/geneseqp-emb1/AA1997.DAT.*
19:	/SIDS1/cgcdata/geneseq/geneseqp-emb1/AA1998.DAT.*
20:	/SIDS1/cgcdata/geneseq/geneseqp-emb1/AA1999.DAT.*
21:	/SIDS1/cgcdata/geneseq/geneseqp-emb1/AA2000.DAT.*
22:	/SIDS1/cgcdata/geneseq/geneseqp-emb1/AA2001.DAT.*
23:	/SIDS1/cgcdata/geneseq/geneseqp-emb1/AA2002.DAT.*
24:	/SIDS1/cgcdata/geneseq/geneseqp-emb1/AA2003.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query $\epsilon$		Length	DB	ID	Description
		Match					
1	76	100.0	16	19	AAW30506		DP-1 transcription
2	76	100.0	19	19	AAW30504		DP-1 transcription
3	76	100.0	28	19	AAW57051		E2F activity inhib
4	76	100.0	28	19	AAW57055		E2F activity inhib
5	76	100.0	37	19	AAW30501		DP-1 transcription
6	68	89.5	19	19	AAW30516		DP-1 transcription
7	68	89.5	19	19	AAW30515		DP-1 transcription
8	67	88.2	30	19	AAW30507		DP-1 transcription
9	64	84.2	14	19	AAW30511		DP-1 transcription

10	64	84.2	19	19	AAW30517	DP-1 transcription
11	58	76.3	15	19	AAW57052	E2F activity inhib
12	56	73.7	83	20	AAW32163	Soybean DP-1 prote
13	55	72.4	56	22	AAW67766	Fragment from a wh
14	46	60.5	20	19	AAW30505	DP-1 transcription
15	46	60.5	96	22	ABG58365	Human liver peptid
16	46	60.5	96	22	ABW42955	Peptide #10461 enc
17	46	60.5	96	22	AAW63859	Human brain expres
18	46	60.5	96	22	AAW76672	Human bone marrow
19	46	60.5	96	22	AAW36777	Peptide #10814 enc
20	46	60.5	96	23	ABG45881	Human peptide enco
21	44	57.9	11	19	AAW30509	DP-1 transcription
22	44	57.9	80	20	AAV32167	Soybean E2F protei
23	42	55.3	9	19	AAW30502	DP-1 transcription
24	41	53.9	35	23	AAU72601	DEF domain consens
25	39	51.3	29	19	AAW57054	E2F activity inhib
26	39	51.3	29	19	AAW57048	E2F activity inhib
27	39	51.3	85	23	AAW78095	Human DNA binding
28	39	51.3	85	24	ABW82986	Repressor protein
29	38	50.0	76	23	AAU78096	Human DNA binding
30	38	50.0	76	24	ABW82987	Repressor activatio
31	37	48.7	24	15	ABW56646	Tyrosine activation
32	37	48.7	28	21	ABY44493	Wheat E2F derived
33	36	47.4	88	23	ABP00754	Human ORFX protein
34	35	46.1	52	22	AAW60343	Human foetal prote
35	35	46.1	63	22	AAU22557	Novel human colon
36	35	46.1	63	22	AAW92530	Human digestive sy
37	35	46.1	67	23	ABP39371	Staphylococcus epi
38	35	46.1	87	22	AAW89362	Human immune/haema
39	35	46.1	90	21	AAW53284	Human colon cancer
40	34	44.7	44	19	AAW79339	Human colon cancer
41	34	44.7	51	21	AAW65070	Staphylococcus aur
42	34	44.7	59	22	AAW60361	Human 5' EST relat
43	34	44.7	60	22	AAW60610	Human prostate can
44	34	44.7	67	22	AAW58362	Protonibacterium
45	34	44.7	80	22	AAU07892	Human brain expres
						Polypeptide sequen

## ALIGNMENTS

RESULT 1
AAW30506
ID AAW30506 standard; Peptide; 16 AA.
XX
AC AC
XX AAW30506;
DT 26-OCT-1998 (first entry)
XX
DE DP-1 transcription factor antagonist
XX
KW DP-1; transcription factor; antagonis
KW cell proliferation; cardiovascular ce
KW surgical stent; therapy.
XX
OS Synthetic.
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT Peptide 5..11
FT FT /note= "Claim 3"
FT Peptide 7..16
FT FT /note= "Claim 3"
XX
PN WO9828334-A1.
XX
PD 02-JUL-1998.
XX
PF 22-DEC-1997; 97WO-GB03506.
XX
PR 20-DEC-1996; 96GB-0026589.
XX
PA (PROL-) PROLIFIX LTD.

DP-1 transcription  
E2F activity inhib  
Soybean DP-1 prote  
Fragment from a wh  
DP-1 transcription  
Human liver peptid  
Peptide 10461 enc  
Human brain expres  
Human bone marrow  
Peptide 10814 enc  
Human peptide enc  
DP-1 transcription  
Soybean E2F protei  
DP-1 transcription  
DEF domain consens  
E2F activity inhib  
E2F activity inhib  
Human DNA binding  
Repressor protein  
Human DNA binding  
Repressor protein  
Tyrosine activation  
Wheat E2F derived  
Human ORF protein  
Human foetal prote  
Human human colon  
Novel human colon  
Human digestive sy  
Staphylococcus epi  
Human immune/haema  
Human colon cancer  
Human colon cancer  
Staphylococcus aur  
Human 5' EST relat  
Human prostate can  
Propionibacterium  
Human brain expres  
Polypeptide sequen

```

XX PA (PROL-) PROLIFIX LTD.
XX PI Bandara LR, La Thangue NB;
XX PI WPI; 1998-377596/32.
XX DR
XX DR Polypeptide fragments of the DP-1 transcription factor - used for
XX PT inducing apoptosis, specifically in tumour and cardiovascular cells,
XX PT e.g. for preventing restenosis
XX PT
XX PS Claim 4; Page 44; 55pp; English.
XX CC
XX CC Peptide H5 comprises amino acid residues 168-183 in the DEF box
XX CC (I) (see AAW30501) of transcription factor DP1. Claimed peptides
XX CC (II) (see AAW30504-07) containing one or both of 2 motifs (see
XX CC AAW30502-03) of the DEF box are capable of antagonising the
XX CC heterodimerisation of a DP protein with an E2F protein. Also
XX CC claimed are variants of these peptides, especially containing
XX CC substitutions of residues corresponding to residues 167, 169, 171
XX CC and 175 of DP-1, fusion proteins (III) comprising (I) or (II) and a
XX CC membrane translocation sequence (see AAW30508), expression vectors
XX CC encoding (I)-(III) and host cells. (I)-(III) are used
XX CC therapeutically to induce apoptosis, specifically in tumour or
XX CC cardiovascular cells, either in vivo or in vitro, e.g. for purging
XX CC bone marrow. Surgical stents comprising (I)-(III) are used to
XX CC treat or prevent restenosis in patients who have undergone
XX CC angioplasty. (I)-(III) function by inactivating the DNA-binding
XX CC activity of DP/E2F heterodimers. They are also used as research
XX CC reagents, as positive controls in assays for identifying
XX CC antagonists of DP-1/E2F dimerisation and as immunoassay agents.
XX CC Also described is the use of sequences antisense to nucleic acids
XX CC encoding (I)-(III) to control DP levels in cells, particularly by
XX CC gene therapy. When formulated with cytotoxic or cytostatic agents,
XX CC (I)-(III) enhance cell killing.
XX SQ Sequence 16 AA;
XX
XX Query Match 100.0%; Score 76; DB 19; Length 16;
XX Best Local Similarity 100.0%; Pred. No. 3e-08; Indels 0; Gaps 0;
XX Matches 16; Conservative 0; Mismatches 0;
XX
XX QY 1 RYVDALNVLMAWNIIS 16
XX Db 1 RYVDALNVLMAWNIIS 16
XX
XX RESULT 2
XX AAW30504
XX ID AAW30504 standard; Peptide; 19 AA.
XX AC AAW30504;
XX XX
XX DT 26-OCT-1998 (first entry)
XX XX
XX DE DP-1 transcription factor antagonist peptide H2.
XX XX
XX KW DP-1; transcription factor; antagonist; E2F protein; apoptosis;
XX KW cell proliferation; cardiovascular cell; restenosis; tumour;
XX KW surgical stent; therapy.
XX XX
XX OS Synthetic.
XX OS Homo sapiens.
XX XX
XX FH Key Location/Qualifiers
XX FT Peptide 9..18
XX FT /note= "Claim 3"
XX XX
XX PN WO9828334-A1.
XX XX
XX PD 02-JUL-1998.
XX XX
XX PF 22-DEC-1997; 97WO-GB03506.
XX XX
XX PR 20-DEC-1996; 96GB-0026589.
XX XX
XX
XX (PROL-) PROLIFIX LTD.
XX Bandara LR, La Thangue NB;
XX WPI; 1998-377596/32.
XX
XX Polypeptide fragments of the DP-1 transcription factor - used for
XX inducing apoptosis, specifically in tumour and cardiovascular cells,
XX e.g. for preventing restenosis
XX
XX Claim 4; Page 44; 55pp; English.
XX
XX Peptide H2 comprises amino acid residues 166-184 in the DEF box
XX (I) (see AAW30501) of transcription factor DP1. Claimed peptides
XX (II) (see AAW30504-07) containing one or both of 2 motifs (see
XX AAW30502-03) of the DEF box are capable of antagonising the
XX heterodimerisation of a DP protein with an E2F protein. Also
XX claimed are variants of these peptides, especially containing
XX substitutions of residues corresponding to residues 167, 169, 171
XX and 175 of DP-1, fusion proteins (III) comprising (I) or (II) and a
XX membrane translocation sequence (see AAW30508), expression vectors
XX encoding (I)-(III) and host cells. (I)-(III) are used
XX therapeutically to induce apoptosis, specifically in tumour or
XX cardiovascular cells, either in vivo or in vitro, e.g. for purging
XX bone marrow. Surgical stents comprising (I)-(III) are used to
XX treat or prevent restenosis in patients who have undergone
XX angioplasty. (I)-(III) function by inactivating the DNA-binding
XX activity of DP/E2F heterodimers. They are also used as research
XX reagents, as positive controls in assays for identifying
XX antagonists of DP-1/E2F dimerisation and as immunoassay agents.
XX Also described is the use of sequences antisense to nucleic acids
XX encoding (I)-(III) to control DP levels in cells, particularly by
XX gene therapy. When formulated with cytotoxic or cytostatic agents,
XX (I)-(III) enhance cell killing.
XX
XX Query Match 100.0%; Score 76; DB 19; Length 19;
XX Best Local Similarity 100.0%; Pred. No. 3.8e-08; Indels 0; Gaps 0;
XX Matches 16; Conservative 0; Mismatches 0;
XX
XX QY 1 RYVDALNVLMAWNIIS 16
XX Db 3 RYVDALNVLMAWNIIS 18
XX
XX RESULT 3
XX AAW57051
XX ID AAW57051 standard; peptide; 28 AA.
XX AC AAW57051;
XX XX
XX DT 28-AUG-1998 (first entry)
XX XX
XX DE E2F activity inhibiting compound Ib-1.
XX KW E2F activity; inhibitor; treatment; tumour; arteriosclerosis.
XX XX
XX OS Synthetic.
XX OS Homo sapiens.
XX XX
XX FH Key Location/Qualifiers
XX FT Modified-site 1
XX FT /note= "N-terminal acetyl"
XX FT Modified-site 28
XX FT /note= "C-terminal amide"
XX XX
XX PN WO9814474-A1.
XX XX
XX PD 09-APR-1998.
XX XX
XX PF 26-SEP-1997; 97WO-JP03442.
XX XX

```



XX 30-SEP-1996; 96JP-0259432.  
 XX (KYOW ) KYOWA HAKKO KOGYO KK.  
 XX Mizukami T, Shibata K, Yamasaki M, Yoshida T;  
 XX WPI; 1998-240020/21.  
 XX E2F activity inhibitors - for treatment and prevention of tumours  
 XX and arteriosclerosis  
 XX Example 3; Page 27; 52pp; Japanese.  
 XX This represents a compound that can inhibit E2F activity. The compound  
 XX is of the formula R1 - A - R2 where R1 is an optionally substituted  
 XX alkanoyl, allyl, hetero-arylcabonyl, alkoxy carbonyl, aryloxy carbonyl,  
 XX hetero-aryloxy carbonyl, or H, R2 is OH, or optionally substituted alkoxy  
 XX or amino, and A is an E2F family dimer forming region or DNA binding  
 XX region, of at least 12 consecutive amino acids. Compounds of this formula  
 XX can be used to inhibit E2F activity, and are useful in the treatment and  
 XX prevention of tumours and arteriosclerosis.  
 XX Sequence 28 AA;  
 XX Query Match 100.0%; Score 76; DB 19; Length 28;  
 XX Best Local Similarity 100.0%; Pred. No. 6.4e-08;  
 XX Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 RYVDALNVLMMNNIIS 16  
 DB 13 RYVDALNVLMMNNIIS 28  
 RESULT 4  
 AAW57055  
 ID AAW57055 standard; peptide; 28 AA.  
 AC AAW57055;  
 XX 28-AUG-1998 (first entry)  
 DE E2F activity inhibiting compound Ib-3.  
 XX E2F activity; inhibitor; treatment; tumour; arteriosclerosis.  
 XX Synthetic.  
 XX Homo sapiens.  
 XX Key Location/Qualifiers  
 FT Modified-site 1  
 FT Modified-site 28 /note= "N-terminal lauroyl"  
 FT Modified-site 28 /note= "C-terminal amide"  
 XX WO9814474-A1.  
 XX 09-APR-1998.  
 XX 26-SEP-1997; 97WO-JP03442.  
 XX 30-SEP-1996; 96JP-0259432.  
 XX (KYOW ) KYOWA HAKKO KOGYO KK.  
 XX Mizukami T, Shibata K, Yamasaki M, Yoshida T;  
 XX WPI; 1998-240020/21.  
 XX E2F activity inhibitors - for treatment and prevention of tumours  
 XX and arteriosclerosis  
 XX Example 7; Page 33; 52pp; Japanese.

XX This represents a compound that can inhibit E2F activity. The compound  
 XX is of the formula R1 - A - R2 where R1 is an optionally substituted  
 XX alkanoyl, allyl, hetero-arylcabonyl, alkoxy carbonyl, aryloxy carbonyl,  
 XX hetero-aryloxy carbonyl, or H, R2 is OH, or optionally substituted alkoxy  
 XX or amino, and A is an E2F family dimer forming region or DNA binding  
 XX region, of at least 12 consecutive amino acids. Compounds of this formula  
 XX can be used to inhibit E2F activity, and are useful in the treatment and  
 XX prevention of tumours and arteriosclerosis.  
 XX Sequence 28 AA;  
 XX Query Match 100.0%; Score 76; DB 19; Length 28;  
 XX Best Local Similarity 100.0%; Pred. No. 6.4e-08;  
 XX Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 RYVDALNVLMMNNIIS 16  
 DB 13 RYVDALNVLMMNNIIS 28  
 RESULT 5  
 AAW30501  
 ID AAW30501 standard; Peptide; 37 AA.  
 XX AAW30501;  
 AC AAW30501;  
 XX 26-OCT-1998 (first entry)  
 DE DP-1 transcription factor peptide H (DEF box).  
 XX DP-1; transcription factor; antagonist; E2F protein; apoptosis;  
 XX cell proliferation; cardiovascular cell; restenosis; tumour;  
 XX surgical stent; therapy.  
 XX Synthetic.  
 XX Homo sapiens.  
 XX WO9828334-A1.  
 XX 02-JUL-1998.  
 XX 22-DEC-1997; 97WO-GH03506.  
 XX 20-DEC-1996; 96GB-0026589.  
 XX (PROL-) PROLIFIX LTD.  
 XX Bandara LR, La Thangue NB;  
 XX WPI; 1998-377596/32.  
 XX Polypeptide fragments of the DP-1 transcription factor - used for  
 XX inducing apoptosis, specifically in tumour and cardiovascular cells,  
 XX e.g. for preventing restenosis  
 XX Claim 1; Page 44; 55pp; English.  
 XX Peptide H (I) comprises residues 163-199, i.e. the DEF box region,  
 XX of transcription factor DP1. Claimed fragments (II) (see AAW30502-07)  
 XX of (I) are capable of antagonising the heterodimerisation of a DP  
 XX protein with an E2F protein. Also claimed are fusion proteins  
 XX (III) comprising (I) or (II) and a membrane translocation sequence  
 XX (see AAW30508), expression vectors encoding (I)-(III) and host cells.  
 XX (I)-(III) are used therapeutically to induce apoptosis,  
 XX specifically in tumour or cardiovascular cells, either in vivo or in  
 XX vitro, e.g. for purging bone marrow. Surgical stents comprising  
 XX (I)-(III) are used to treat or prevent restenosis in patients who  
 XX have undergone angioplasty. (I)-(III) function by inactivating  
 XX the DNA-binding activity of DP/E2F heterodimers. They are also  
 XX used as research reagents, as positive controls in assays for  
 XX identifying antagonists of DP-1/E2F dimerisation and as immunoassay  
 XX agents. Also described is the use of sequences antisense to

CC nucleic acids encoding (I)-(III) to control DP levels in cells,  
CC particularly by gene therapy. When formulated with cytotoxic  
CC or cytostatic agents, (I)-(III) enhance cell killing.

SQ Sequence 37 AA;

Query Match 100.0%; Score 76; DB 19; Length 37;  
Best Local Similarity 100.0%; Pred. No. 9.3e-08;  
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 RYVDALNVLMMNNIIS 16  
DB 6 RYVDALNVLMMNNIIS 21

RESULT 6  
AAW30516  
ID AAW30516 standard; Peptide; 19 AA.

XX AC AAW30516;  
XX DT 26-OCT-1998 (first entry)  
XX DE DP-1 transcription factor antagonist peptide H2mt2.  
XX KW DP-1; transcription factor; E2F protein; apoptosis;  
XX KW cell proliferation; cardiovascular cell; restenosis; tumour;  
XX KW surgical stent; therapy.

XX OS Synthetic.  
XX OS Homo sapiens.

XX FH Key Location/Qualifiers  
XX FT Misc-difference 2  
XX FT Misc-difference 6 /note= "R167A mutation"  
XX FT Misc-difference 6 /note= "D171A mutation"

XX PN WO9828334-A1.  
XX PD 02-JUL-1998.  
XX PF 22-DEC-1997; 97WO-GB03506.  
XX PR 20-DEC-1996; 96GB-0026589.

XX PA (PROL-) PROLIFIX LTD.  
XX PI Bandara LR, La Thangue NB;  
XX DR WPI; 1998-377596/32.

XX PT Polypeptide fragments of the DP-1 transcription factor - used for  
XX PT inducing apoptosis, specifically in tumour and cardiovascular cells,  
XX PT e.g. for preventing restenosis  
XX PS Example D; Page 26; 55pp; English.

XX CC Peptide H2mt2 is based on peptide H2 (see AAW30504) from the DEF box  
XX CC (see AAW30501) of transcription factor DP1. In H2mt2, amino acid  
XX CC residues of H2 that correspond to DP1 residues Arg167 and Asp171  
XX CC are substituted by Ala residues. H2 is an antagonist of the  
XX CC heterodimerisation of DP1 with E2F. H2mt2 retains some, but not  
XX CC all, of this antagonistic activity. H2 and other claimed peptides  
XX CC (see AAW30504-07) from the DEF box region of DP1 can be used to  
XX CC induce apoptosis, specifically in tumour and cardiovascular cells,  
XX CC e.g. for the prevention of restenosis.

SQ Sequence 19 AA;

Query Match 89.5%; Score 68; DB 19; Length 19;  
Best Local Similarity 93.8%; Pred. No. 1.5e-06;  
Matches 15; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 RYVDALNVLMMNNIIS 16  
DB 3 RYVDALNVLMMNNIIS 18

RESULT 7  
AAW30515  
ID AAW30515 standard; Peptide; 19 AA.

XX AC AAW30515;  
XX DT 26-OCT-1998 (first entry)  
XX DE DP-1 transcription factor peptide H2mt1.  
XX KW DP-1; transcription factor; antagonist; E2F protein; apoptosis;  
XX KW cell proliferation; cardiovascular cell; restenosis; tumour;  
XX KW surgical stent; therapy.

XX OS Synthetic.  
XX OS Homo sapiens.

XX FH Key Location/Qualifiers  
XX FT Misc-difference 4 /note= "V169A mutation"  
XX FT Misc-difference 10 /note= "V175A mutation"

XX PN WO9828334-A1.  
XX PD 02-JUL-1998.  
XX PF 22-DEC-1997; 97WO-GB03506.  
XX PR 20-DEC-1996; 96GB-0026589.

XX PA (PROL-) PROLIFIX LTD.  
XX PI Bandara LR, La Thangue NB;  
XX DR WPI; 1998-377596/32.

XX PT Polypeptide fragments of the DP-1 transcription factor - used for  
XX PT inducing apoptosis, specifically in tumour and cardiovascular cells,  
XX PT e.g. for preventing restenosis  
XX PS Example D; Page 26; 55pp; English.

XX CC Peptide H2mt1 is based on peptide H2 (see AAW30504) from the DEF box  
XX CC (see AAW30501) of transcription factor DP1. The H2mt1 peptide, in  
XX CC which H2 residues corresponding to DP1 residues Val169 and Val175  
XX CC are substituted by Ala residues, behaves in a similar fashion to  
XX CC the wild-type H2 peptide in its ability to inactivate E2F site DNA  
XX CC binding activity in D9 EC cell extracts. H2 is an antagonist of the  
XX CC heterodimerisation of DP1 with E2F. Thus, the Val-169 and Val-175  
XX CC residues of H2 play little role in this activity. H2 and other  
XX CC claimed peptides (see AAW30504-07) from the DEF box region of DP1 can  
XX CC be used to induce apoptosis, specifically in tumour and  
XX CC cardiovascular cells, e.g. for the prevention of restenosis.

SQ Sequence 19 AA;

Query Match 89.5%; Score 68; DB 19; Length 19;  
Best Local Similarity 87.5%; Pred. No. 1.5e-06;  
Matches 14; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 1 RYVDALNVLMMNNIIS 16  
DB 3 RYVDALNVLMMNNIIS 18

RESULT 8

AAW30507  
 ID AAW30507 standard; Peptide; 30 AA.  
 AC AAW30507;  
 XX  
 DT 26-OCT-1998 (first entry)  
 XX  
 DE DP-1 transcription factor antagonist peptide H7.  
 XX  
 KW DP-1; transcription factor; antagonist; E2F protein; apoptosis;  
 KW cell proliferation; cardiovascular cell; restenosis; tumour;  
 KW surgical stent; therapy.  
 XX  
 OS Synthetic.  
 OS Homo sapiens.  
 XX  
 FH Key Location/Qualifiers  
 FT Peptide 3..9  
 FT /note= "Claim 3"  
 FT 5..15  
 FT /note= "Claim 3"  
 FT  
 PN WO9828334-A1.  
 XX  
 PD 02-JUL-1998.  
 XX  
 PF 22-DEC-1997; 97WO-GB03506.  
 XX  
 PR 20-DEC-1996; 96GB-0026589.  
 XX  
 PA (PROL-) PROLIFIX LTD.  
 XX  
 PI Bandara LR, La Thangue NB;  
 XX  
 DR WPI; 1998-377596/32.  
 XX  
 PT Polypeptide fragments of the DP-1 transcription factor - used for  
 PT inducing apoptosis, specifically in tumour and cardiovascular cells,  
 PT e.g. for preventing restenosis  
 XX  
 PS Claim 4; Page 44; 55pp; English.  
 XX  
 CC Peptide H7 comprises amino acid residues 170-199 in the DEF box  
 CC (I) (see AAW30501) of transcription factor DP1. Claimed peptides  
 CC (II) (see AAW30504-07) containing one or both of 2 motifs (see  
 CC AAW30502-03) of the DEF box are capable of antagonising the  
 CC heterodimerisation of a DP protein with an E2F protein. Also  
 CC claimed are variants of these peptides, especially containing  
 CC substitutions of residues corresponding to residues 167, 169, 171  
 CC and 175 of DP-1, fusion proteins (III) comprising (I) or (II) and a  
 CC membrane translocation sequence (see AAW30508), expression vectors  
 CC encoding (I)-(III) and host cells. (I)-(III) are used  
 CC therapeutically to induce apoptosis, specifically in tumour or  
 CC cardiovascular cells, either in vivo or in vitro, e.g. for purging  
 CC bone marrow. Surgical stents comprising (I)-(III) are used to  
 CC treat or prevent restenosis in patients who have undergone  
 CC angioplasty. (I)-(III) function by inactivating the DNA-binding  
 CC activity of DP/E2F heterodimers. They are also used as research  
 CC reagents, as positive controls in assays for identifying  
 CC antagonists of DP-1/E2F dimerisation and as immunoassay agents.  
 CC Also described is the use of sequences antisense to nucleic acids  
 CC encoding (I)-(III) to control DP levels in cells, particularly by  
 CC gene therapy. When formulated with cytotoxic or cytostatic agents,  
 CC (I)-(III) enhance cell killing.  
 XX  
 SQ Sequence 30 AA;  
 Query Match 88.2%; Score 67; DB 19; Length 30;  
 Best Local Similarity 100.0%; Pred. No. 4.2e-06;  
 Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 3 YDALNVLMMNIIIS 16  
 |||||

Db 1 YDALNVLMMNIIIS 14  
 RESULT 9  
 AAW30511  
 ID AAW30511 standard; Peptide; 14 AA.  
 XX  
 AC AAW30511;  
 XX  
 DT 26-OCT-1998 (first entry)  
 XX  
 DE DP-1 transcription factor peptide H6.  
 XX  
 KW DP-1; transcription factor; antagonist; E2F protein; apoptosis;  
 KW cell proliferation; cardiovascular cell; restenosis; tumour;  
 KW surgical stent; therapy.  
 XX  
 OS Synthetic.  
 OS Homo sapiens.  
 XX  
 PN WO9828334-A1.  
 XX  
 PD 02-JUL-1998.  
 XX  
 PF 22-DEC-1997; 97WO-GB03506.  
 XX  
 PR 20-DEC-1996; 96GB-0026589.  
 XX  
 PA (PROL-) PROLIFIX LTD.  
 XX  
 PI Bandara LR, La Thangue NB;  
 XX  
 DR WPI; 1998-377596/32.  
 XX  
 PT Polypeptide fragments of the DP-1 transcription factor - used for  
 PT inducing apoptosis, specifically in tumour and cardiovascular cells,  
 PT e.g. for preventing restenosis  
 XX  
 PS Example C; Page 41; 55pp; English.  
 XX  
 CC Peptide H6 comprises amino acid residues 167-180 in the DEF box  
 CC region (see AAW30501) of transcription factor DP1. Unlike claimed  
 CC peptides (see AAW30504-07) that contain one or both of 2 motifs (see  
 CC AAW30502-03) of the DP1 DEF box, peptide H6 is not capable of  
 CC antagonising the heterodimerisation of a DP protein with an E2F  
 CC protein. The claimed peptides, their variants and fusion proteins  
 CC can be used to induce apoptosis, specifically in tumour and  
 CC cardiovascular cells, e.g. to prevent restenosis.  
 XX  
 SQ Sequence 14 AA;  
 Query Match 84.2%; Score 64; DB 19; Length 14;  
 Best Local Similarity 100.0%; Pred. No. 6e-06;  
 Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 RYVDALNVLMMN 13  
 |||||  
 Db 2 RYVDALNVLMMN 14  
 RESULT 10  
 AAW30517  
 ID AAW30517 standard; Peptide; 19 AA.  
 XX  
 AC AAW30517;  
 XX  
 DT 26-OCT-1998 (first entry)  
 XX  
 DE DP-1 transcription factor peptide H2mt3.  
 XX  
 KW DP-1; transcription factor; antagonist; E2F protein; apoptosis;  
 KW cell proliferation; cardiovascular cell; restenosis; tumour;  
 KW surgical stent; therapy.

XX Synthetic.  
OS Homo sapiens.  
XX  
XX Key Location/Qualifiers  
FH Misc-difference 8 /note= "L173R mutation"  
FT Misc-difference 11  
FT Misc-difference 11 /note= "L176R mutation"  
XX  
XX WO9828334-A1.  
XX  
XX 02-JUL-1998.  
XX  
XX 22-DEC-1997; 97WO-GB03506.  
XX  
XX 20-DEC-1996; 96GB-0026589.  
XX (PROL-) PROLIFIX LTD.  
XX Bandara LR, La Thangue NB;  
XX WPI; 1998-377596/32.  
XX Polypeptide fragments of the DP-1 transcription factor - used for  
FT inducing apoptosis, specifically in tumour and cardiovascular cells,  
FT e.g. for preventing restenosis  
XX  
XX Example D; Page 26; 55pp; English.  
XX Peptide H2mt3 is based on peptide H2 (see AAW30504) from the DEF box  
CC (see AAW30501) of transcription factor Dp1. In H2mt3, amino acid  
CC residues of H2 that correspond to Dp1 residues Leu173 and Leu176  
CC are substituted by Arg residues. H2 is an antagonist of the  
CC heterodimerisation of Dp1 with E2F. H2mt3 has none of the  
CC antagonistic activity of H2. H2 and other claimed peptides (see  
CC AAW30504-07) from the DEF box region of Dp1 can be used to induce  
CC apoptosis, specifically in tumour and cardiovascular cells, e.g.  
CC for the prevention of restenosis.  
XX  
XX Sequence 19 AA;  
Query Match 84.2%; Score 64; DB 19; Length 19;  
Best Local Similarity 87.5%; Pred. No. 8.9e-06;  
Matches 14; Conservative 0; Mismatches 2; Indels 0; Gaps 0;  
QY 1 RYVDALNVLMMNIIIS 16  
DB 3 RYVDARNVRMMNIIIS 18  
RESULT 11  
AAW57052  
ID AAW57052 standard; peptide; 15 AA.  
XX  
XX AAW57052;  
XX  
XX 28-AUG-1998 (first entry)  
DE E2F activity inhibiting compound Ib-2.  
XX  
XX E2F activity; inhibitor; treatment; tumour; arteriosclerosis.  
OS Synthetic.  
OS Homo sapiens.  
XX Key Location/Qualifiers  
FH Modified-site 1 /note= "N-terminal acetyl"  
FT Modified-site 15  
FT Modified-site 15 /note= "C-terminal amide"  
XX  
XX WO9814474-A1.  
PN

XX 09-APR-1998.  
XX  
XX 26-SEP-1997; 97WO-JP03442.  
XX  
XX 30-SEP-1996; 96JP-0259432.  
XX (KYOW ) KYOWA HAKKO KOGYO KK.  
XX Mizukami T, Shibata K, Yamasaki M, Yoshida T;  
PI WPI; 1998-240020/21.  
XX  
XX E2F activity inhibitors - for treatment and prevention of tumours  
PT and arteriosclerosis  
XX  
XX Example 4; Page 28; 52pp; Japanese.  
XX This represents a compound that can inhibit E2F activity. The compound  
CC is of the formula R1 - A - R2 where R1 is an optionally substituted  
CC alkanoyl, allyl, hetero-arylcarbonyl, alkoxyarbonyl, aryloxyarbonyl,  
CC hetero-aryloxyarbonyl, or H, R2 is OH, or optionally substituted alkoxy  
CC or amino, and A is an E2F family dimer forming region or DNA binding  
CC region, of at least 12 consecutive amino acids. Compounds of this formula  
CC can be used to inhibit E2F activity, and are useful in the treatment and  
CC prevention of tumours and arteriosclerosis.  
XX  
XX Sequence 15 AA;  
Query Match 76.3%; Score 58; DB 19; Length 15;  
Best Local Similarity 100.0%; Pred. No. 0.0001;  
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 RYVDALNVLMM 12  
DB 4 RYVDALNVLMM 15  
RESULT 12  
AAV32163  
ID AAV32163 standard; Protein; 83 AA.  
XX  
XX AAV32163;  
XX  
XX 01-FEB-2000 (first entry)  
DT  
XX Soybean DP-1 protein fragment.  
DE  
XX DP-1; soybean; cell cycle regulatory protein;  
KW transcription factor; herbicide.  
XX  
XX Glycine max.  
XX  
XX Key Location/Qualifiers  
FH Misc-difference 10 /note= "encoded by GNC"  
FT Misc-difference 25 /note= "encoded by ANT"  
FT Misc-difference 26 /note= "encoded by GNC"  
FT Misc-difference 34 /note= "encoded by GNC"  
FT Misc-difference 35 /note= "encoded by GGN"  
FT Misc-difference 49 /note= "encoded by CNT"  
FT Misc-difference 49 /note= "encoded by NAN"  
FT Misc-difference 54 /note= "encoded by ANG"  
FT Misc-difference 55 /note= "encoded by NAT"  
FT Misc-difference 59 /note= "encoded by NAG"  
FT Misc-difference 63

FT Misc-difference 69 /note= "encoded by NAT"  
 FT FT /note= "encoded by GNG"  
 FT Misc-difference 71  
 FT FT /note= "encoded by NTC"  
 FT Misc-difference 80  
 FT FT /note= "encoded by CNA"

XX WO9953075-A2.  
 PN  
 XX 21-OCT-1999.  
 PD  
 XX 08-APR-1999; 99WO-US07638.  
 PF  
 XX 09-APR-1998; 98US-0081132.  
 PR  
 XX (DUPO ) DU PONT DE NEMOURS & CO E I.  
 PA  
 XX Klein TM, Morakinyo LO, Odell JT, Sakai H;  
 PI  
 XX WPI; 1999-633830/54.  
 DR N-PSDB; AA234579.  
 DR

XX Plant-derived cell cycle regulatory proteins -  
 PT  
 XX Claim 10; Page 41; 44pp; English.

XX This sequence represents 42% of the middle region of soybean cell  
 CC cycle regulatory protein DP-1, as deduced from an isolated  
 CC cDNA clone (see AA234579). The invention relates to nucleic acid  
 CC fragments (see AA234575-83) encoding plant CDC-16, DP-1, DP-2 and  
 CC E2F cell cycle regulatory proteins (see AA232159-67). It also  
 CC relates to the construction of a chimeric gene encoding all or a  
 CC portion of the cell cycle regulatory protein, in sense or antisense  
 CC orientation, where expression of the chimeric gene results in  
 CC production of altered levels of the cell cycle regulatory protein in  
 CC a transformed host cell. The nucleic acids and proteins may be  
 CC used to facilitate studies of cell cycle regulation in plants,  
 CC provide genetic tools to enhance cell growth in tissue culture,  
 CC increase gene transfer efficiency and provide more stable  
 CC transformations. The proteins may also provide targets to  
 CC facilitate design and/or identification of cell cycle regulatory  
 CC proteins that may be useful as herbicides.

XX SQ Sequence 83 AA;

Query Match 73.7%; Score 56; DB 20; Length 83;  
 Best Local Similarity 81.2%; Pred. No. 0.0024;  
 Matches 13; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 RYVDALNVLMMNIIIS 16  
 | ||||| |||||  
 Db 9 RXDDALNVLMMNIIIS 24

RESULT 13  
 AAB67766

ID AAB67766 standard; peptide; 56 AA.

XX AAB67766;

XX 11-JUN-2001 (first entry)

DE Fragment from a wheat E2F-dimerisation partner (DP) protein.

XX E2F-dimerisation partner; DP protein; E2F transcription factor;  
 KW G1 phase; S phase; cell cycle; retinoblastoma protein;  
 KW alter cell proliferation.

XX Triticum monococcum.

XX WO200121644-A2.

XX

PD 29-MAR-2001.  
 XX  
 XX 25-SEP-2000; 2000WO-EP09325.  
 PF  
 XX 24-SEP-1999; 99ES-0002127.  
 PR 11-NOV-1999; 99ES-0002474.  
 XX  
 XX (CNSJ ) CONSEJO SUPERIOR INVESTIGACIONES CIENTIF.  
 PA  
 XX Gutierrez-Armenta C, Ramirez-Parra E;  
 PI WPI; 2001-257972/26.  
 DR N-PSDB; AAF80148.  
 DR  
 XX New isolated, enriched, cell free and/or recombinant nucleic acid  
 PT useful for e.g. altering cell proliferation characteristic such as to  
 PT alter plant cell, organ or tissue size -  
 PT  
 XX Claim 10; Page 50; 77pp; English.  
 PS  
 XX AAB67764-68 represent fragments of a wheat E2F-dimerisation partner  
 CC (DP) protein. The protein acts as a plant E2F transcription factor.  
 CC E2F and DP are two proteins that hetero-dimerise to form an active  
 CC transcription factor that regulates G1 to S phase of the cell cycle,  
 CC and later, the expression of genes required for S-phase progression.  
 CC E2F and retinoblastoma protein also interact as a hetero-dimer in  
 CC cells to suppress certain genes. This repression involves binding of  
 CC the retinoblastoma protein to the E2F-DP dimer that is in turn bound  
 CC to sites on DNA through the E2F DNA binding domain. DP proteins can  
 CC be modulated to alter plant cell, organ or tissue shape, and  
 CC particularly to alter cell proliferation characteristic such as to  
 CC alter plant cell, organ or tissue size.

XX SQ Sequence 56 AA;

Query Match 72.4%; Score 55; DB 22; Length 56;  
 Best Local Similarity 62.5%; Pred. No. 0.0023;  
 Matches 10; Conservative 4; Mismatches 2; Indels 0; Gaps 0;

QY 1 RYVDALNVLMMNIIIS 16  
 ||||| |||||  
 Db 16 RYVDALNVLMMNIIIS 31

RESULT 14  
 AAW30505

ID AAW30505 standard; Peptide; 20 AA.

XX AAW30505;

XX 26-OCT-1998 (first entry)

DE DP-1 transcription factor antagonist peptide H3.

XX DP-1; transcription factor; antagonist; E2F protein; apoptosis;  
 KW cell proliferation; cardiovascular cell; restenosis; tumour;  
 KW surgical stent; therapy.

XX Synthetic.

XX Homo sapiens.

XX Key Location/Qualifiers

FT Peptide 1..10

FT /note= "Claim 3"

XX WO9828334-A1.

XX 02-JUL-1998.

XX 22-DEC-1997; 97WO-GB03506.

XX 20-DEC-1996; 96GB-0026589.

XX

PA	(PROL-) PROLIFIX LTD.
XX	Bandara LR, La Thangue NB;
PI	WFI; 1998-377596/32.
XX	Polypeptide fragments of the DP-1 transcription factor - used for inducing apoptosis, specifically in tumour and cardiovascular cells, e.g. for preventing restenosis
XX	Claim 4; Page 44; 55pp; English.
DR	Peptide H3 comprises amino acid residues 174-193 in the DEF box (I) (see AAW30501) of transcription factor DPl. Claimed peptides (II) (see AAW30504-07) containing one or both of 2 motifs (see AAW30502-03) of the DEF box are capable of antagonising the heterodimerisation of a DP protein with an E2F protein. Also claimed are variants of these peptides, especially containing substitutions of residues corresponding to residues 167, 169, 171 and 175 of DP-1, fusion proteins (III) comprising (I) or (II) and a membrane translocation sequence (see AAW30508), expression vectors encoding (I)-(III) and host cells. (I)-(III) are used therapeutically to induce apoptosis, specifically in tumour or cardiovascular cells, either in vivo or in vitro, e.g. for purging bone marrow. Surgical stents comprising (I)-(III) are used to treat or prevent restenosis in patients who have undergone angioplasty. (I)-(III) function by inactivating the DNA-binding activity of DP/E2F heterodimers. They are also used as research reagents, as positive controls in assays for identifying antagonists of DP-1/E2F dimerisation and as immunoassay agents. Also described is the use of sequences antisense to nucleic acids encoding (I)-(III) to control DP levels in cells, particularly by gene therapy. When formulated with cytotoxic or cytostatic agents, (I)-(III) enhance cell killing.
XX	Sequence 20 AA;
XX	Query Match 60.5%; Score 46; DB 19; Length 20;
XX	Best Local Similarity 100.0%; Pred. No. 0.034; Indels 0; Gaps 0;
XX	Matches 10; Conservative 0; Mismatches 0;
OY	7 NVLMANNIIS 16
DB	1 NVLMANNIIS 10
XX	RESULT 15
XX	ABG58365
ID	ABG58365 standard; Peptide; 96 AA.
XX	AC ABG58365;
DT	25-FEB-2003 (first entry)
DE	Human liver peptide, SEQ ID NO 37013.
XX	Human; liver; cirrhosis; hyperlipoproteinaemia; hyperlipidaemia; hypercholesterolaemia; coronary heart disease.
KW	Homo sapiens.
OS	WO200157273-A2.
PX	09-AUG-2001.
PN	30-JAN-2001; 2001WO-US00664.
PD	04-FEB-2000; 2000US-0180312.
XX	26-MAY-2000; 2000US-0207456.
PR	30-JUN-2000; 2000US-0608408.
PR	03-AUG-2000; 2000US-0632366.
PR	21-SEP-2000; 2000US-0234687.
PR	27-SEP-2000; 2000US-0234359.

PT Human genome-derived single exon nucleic acid probes useful for  
 PT analyzing gene expression in human fetal liver -  
 XX  
 PS Claim 27; SEQ ID NO 35590; 639pp + sequence listing; English.  
 CC The invention relates to a single exon nucleic acid probe for  
 CC measuring human gene expression in a sample derived from human foetal  
 CC liver. The single exon nucleic acid probes may be used for predicting,  
 CC measuring and displaying gene expression in samples derived from human  
 CC fetal liver. The present sequence is a peptide encoded by a single exon  
 CC nucleic acid probe of the invention.  
 CC Note: The sequence data for this patent did not form part of the  
 CC printed specification, but was obtained in electronic format directly  
 CC from WIPO at ftp.wipo.int/pub/published\_pct\_sequences.  
 XX  
 SQ Sequence 96 AA;

Query Match 60.5%; Score 46; DB 22; Length 96;  
 Best Local Similarity 43.8%; Pred. No. 0.28;  
 Matches 7; Conservative 7; Mismatches 2; Indels 0; Gaps 0;

QY 1 RYVDALNVLMMAMNIS 16  
 :||:||||:|  
 Db 5 RYDINVNVLESLSLVS 20

RESULT 17  
 AAM63859  
 ID AAM63859 standard; Protein; 96 AA.  
 XX  
 AC AAM63859;  
 XX  
 XX 05-NOV-2001 (first entry)  
 XX  
 XX Human brain expressed single exon probe encoded protein SEQ ID NO: 35964.  
 DE  
 XX Human; brain expressed exon; gene expression analysis; probe;  
 XX microarray; Alzheimer's disease; multiple sclerosis; schizophrenia;  
 KW epilepsy; cancer.  
 KW  
 XX Homo sapiens.  
 OS  
 XX WO200157275-A2.  
 PN  
 XX 09-AUG-2001.  
 PD  
 XX 30-JAN-2001; 2001WO-US00667.  
 PF  
 XX 04-FEB-2000; 2000US-0180312.  
 PR 26-MAY-2000; 2000US-0207456.  
 PR 30-JUN-2000; 2000US-0608408.  
 PR 03-AUG-2000; 2000US-0632366.  
 PR 21-SEP-2000; 2000US-0234687.  
 PR 27-SEP-2000; 2000US-0236359.  
 PR 04-OCT-2000; 2000GB-0024263.  
 XX  
 PA (MOLE-) MOLECULAR DYNAMICS INC.  
 PI Penn SG, Hanzel DK, Chen W, Rank DR;  
 XX WPI; 2001-483446/52.  
 XX  
 XX Single exon nucleic acid probes for analyzing gene expression in human  
 XX brains -  
 XX  
 XX Example 4; SEQ ID NO: 35964; 650pp + Sequence Listing; English.  
 XX  
 XX The present invention provides a number of single exon nucleic acid  
 XX probes which are derived from genomic sequences expressed in the human  
 XX brain. They can be used to measure gene expression in brain cell samples,  
 XX which may enable the diagnosis and improved treatment of nervous system  
 XX diseases such as Alzheimer's disease, multiple sclerosis, schizophrenia,  
 XX epilepsy and cancers. The present sequence is a protein encoded by one of

CC the probes of the invention.  
 XX  
 SQ Sequence 96 AA;

Query Match 60.5%; Score 46; DB 22; Length 96;  
 Best Local Similarity 43.8%; Pred. No. 0.28;  
 Matches 7; Conservative 7; Mismatches 2; Indels 0; Gaps 0;

QY 1 RYVDALNVLMMAMNIS 16  
 :||:||||:|  
 Db 5 RYDINVNVLESLSLVS 20

RESULT 18  
 AAM76672  
 ID AAM76672 standard; Protein; 96 AA.  
 XX  
 AC AAM76672;  
 XX  
 XX 06-NOV-2001 (first entry)  
 XX  
 XX Human bone marrow expressed probe encoded protein SEQ ID NO: 36978.  
 DE  
 XX Human; bone marrow expressed exon; gene expression analysis; probe;  
 KW microarray; cancer; leukaemia; lymphoma; myeloma.  
 KW  
 XX Homo sapiens.  
 OS  
 XX WO200157276-A2.  
 PN  
 XX 09-AUG-2001.  
 PD  
 XX 30-JAN-2001; 2001WO-US00668.  
 PF  
 XX 04-FEB-2000; 2000US-0180312.  
 PR 26-MAY-2000; 2000US-0207456.  
 PR 30-JUN-2000; 2000US-0608408.  
 PR 03-AUG-2000; 2000US-0632366.  
 PR 21-SEP-2000; 2000US-0234687.  
 PR 27-SEP-2000; 2000US-0236359.  
 PR 04-OCT-2000; 2000GB-0024263.  
 XX  
 PA (MOLE-) MOLECULAR DYNAMICS INC.  
 PI Penn SG, Hanzel DK, Chen W, Rank DR;  
 XX WPI; 2001-488900/53.  
 XX  
 XX Human genome-derived single exon nucleic acid probes useful for  
 XX analyzing gene expression in human bone marrow -  
 XX  
 XX Example 4; SEQ ID NO: 36978; 658pp + Sequence Listing; English.  
 XX  
 XX The present invention provides a number of single exon nucleic acid  
 XX probes which are derived from genomic sequences expressed in the human  
 XX bone marrow. They can be used to measure gene expression in bone marrow  
 XX samples, which may enable the improved diagnosis and treatment of cancers  
 XX such as lymphoma, leukaemia and myeloma. The present sequence is a  
 XX protein encoded by one of the probes of the invention.  
 XX  
 SQ Sequence 96 AA;

Query Match 60.5%; Score 46; DB 22; Length 96;  
 Best Local Similarity 43.8%; Pred. No. 0.28;  
 Matches 7; Conservative 7; Mismatches 2; Indels 0; Gaps 0;

QY 1 RYVDALNVLMMAMNIS 16  
 :||:||||:|  
 Db 5 RYDINVNVLESLSLVS 20

RESULT 19  
 AAM36777

primary ciliary dyskinesia; pulmonary hypertension; hyaline membrane disease.

Homo sapiens.

WO200186003-A2.

15-NOV-2001.

30-JAN-2001; 2001WO-US00665.

04-FEB-2000; 2000US-180312P.

26-MAY-2000; 2000US-207456P.

30-JUN-2000; 2000US-0608408.

03-AUG-2000; 2000US-0632366.

21-SEP-2000; 2000US-234687P.

27-SEP-2000; 2000US-236359P.

04-OCT-2000; 2000GB-0024263.

(MOLE-) MOLECULAR DYNAMICS INC.

Penn SG, Hanzel DK, Chen W, Rank DR;

WPI; 2002-114183/15.

Spatially-addressable set of single exon nucleic acid probes, used to measure gene expression in human lung samples -

Claim 27; SEQ ID No 35546; 634pp; English.

The invention relates to a spatially-addressable set of single exon nucleic acid probes for measuring gene expression in a sample derived from human lung comprising single exon nucleic acid probes having one of 12614 nucleic acid sequences mentioned in the specification, or their complements or the 12387 open reading frames derived from the 12614 probes. Also included are a microarray comprising the novel set of probes; the novel set of probes which hybridise at high stringency to a nucleic acid expressed in the human lung; measuring gene expression in a sample derived from human lung, comprising (a) contacting the array with a collection of detectably labeled nucleic acids derived from human lung mRNA, and (b) measuring the labeled nucleic acids bound to each probe of the array; identifying exons in a eukaryotic genome, comprising (a) algorithmically predicting at least one exon from genomic sequences of the eukaryote; and (b) detecting specific hybridisation of detectably labeled nucleic acids from eukaryote lung mRNA, to a single exon probe, having a fragment identical to the predicted exon, the probe is included in the above mentioned microarray; assigning exons to a single gene, comprising (a) identifying exons from genomic sequence by the method above and (b) measuring the expression of each of the exons in several tissues and/or cell types using hybridisation to a single exon microarrays having a probe with the exon, where a common pattern of expression of the exons in the tissues and/or cell types indicates that the exons should be assigned to a single gene; a peptide comprising one of 12011 sequences, mentioned in the specification, or encoded by the probes/open reading frames (ORF). The probes are used for gene expression analysis, and for identifying exons in a gene, particularly using human lung derived mRNA and for the study of lung diseases such as asthma, lung cancer, chronic obstructive pulmonary disease (COPD), interstitial lung disease (ILD), familial idiopathic pulmonary fibrosis, neurofibromatosis, tuberous sclerosis, Gaucher's disease, Niemann-Pick disease, Hermansky-Pudlak syndrome, sarcoidosis, pulmonary haemosiderosis, pulmonary histiocytosis, lymphangioleiomyomatosis, pulmonary alveolar proteinosis, Karagener syndrome, fibrocystic pulmonary dysplasia, primary ciliary dyskinesia, pulmonary hypertension and hyaline membrane disease. The present sequence is a peptide/protein encoded by a single exon probe of the invention.

Note: The sequence data for this patent did not form part of the printed specification, but was obtained in electronic format directly from WIPO at ftp.wipo.int/pub/published\_pct\_sequences.

Sequence 96 AA;



```
Query Match      60.5%; Score 46; DB 23; Length 96;
Best Local Similarity 43.8%; Pred. No. 0.28;
Matches 7; Conservative 7; Mismatches 2; Indels 0; Gaps 0;

QY 1 RYVDALNVLMMNNIIS 16
   |:|:|:|:|:|:|:|
Db 5 RIYDIVNVLESILHS 20

RESULT 21
AAW30509
ID AAW30509 standard; Peptide; 11 AA.
XX
AC AAW30509;
XX
DT 26-OCT-1998 (first entry)
XX
DE DP-1 transcription factor peptide H1.
XX
KW DP-1; transcription factor; antagonist; E2F protein; apoptosis;
KW cell proliferation; cardiovascular cell; restenosis; tumour;
KW surgical stent; therapy.
XX
OS Synthetic.
OS Homo sapiens.
XX
PN WO9828334-A1.
XX
PD 02-JUL-1998.
XX
PF 22-DEC-1997; 97WO-GB03506.
XX
PR 20-DEC-1996; 96GB-0026589.
XX
PA (PROL-) PROLIFIX LTD.
XX
PI Bandara LR, La Thangue NB;
XX
DR WPI; 1998-377596/32.
XX
PT Polypeptide fragments of the DP-1 transcription factor - used for
PT inducing apoptosis, specifically in tumour and cardiovascular cells,
PT e.g. for preventing restenosis
XX
PS Example C; Page 41; 55pp; English.
XX
CC Peptide H1 comprises amino acid residues 166-176 in the DEF box
CC region (see AAW30501) of transcription factor Dp1. Unlike claimed
CC peptides (see AAW30504-07) that contain one or both of 2 motifs (see
CC AAW30502-03) of the Dp1 DEF box, peptide H1 is not capable of
CC antagonising the heterodimerisation of a DP protein with an E2F
CC protein. The claimed peptides, their variants and fusion proteins
CC can be used to induce apoptosis, specifically in tumour and
CC cardiovascular cells, e.g. to prevent restenosis.
XX
SQ Sequence 11 AA;

Query Match      57.9%; Score 44; DB 19; Length 11;
Best Local Similarity 100.0%; Pred. No. 0.038;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RYVDALNVL 9
   |:|:|:|:|:|
Db 3 RYVDALNVL 11

RESULT 22
AAV32167
ID AAY32167 standard; Protein; 80 AA.
XX
AC AAY32167;
XX
DT 01-FEB-2000 (first entry)
```

```
XX Soybean E2F protein fragment.
DE
XX E2F; soybean; cell cycle regulatory protein; transcription factor;
KW herbicide.
KW
XX Glycine max.
OS
XX WO9953075-A2.
PN
XX 21-OCT-1999.
PD
XX 08-APR-1999; 99WO-US07638.
PF
XX 09-APR-1998; 98US-0081132.
PR
XX (DUPO ) DU PONT DE NEMOURS & CO E I.
PA
XX Klein TM, Morakinyo LO, Odell JT, Sakai H;
PI WPI; 1999-633830/54.
XX N-PSDB; AAZ34583.
DR
XX Plant-derived cell cycle regulatory proteins
PT
XX Claim 20; Page 44-45; 44pp; English.
XX
CC This sequence represents 10% of the middle region of soybean
CC cell cycle regulatory protein E2F, as deduced from an isolated
CC cDNA clone (see AAZ34583). The invention relates to nucleic acid
CC fragments (see AAZ34575-83) encoding plant CDC-16, DP-1, DP-2 and
CC E2F cell cycle regulatory proteins (see AAY32159-67). It also
CC relates to the construction of a chimeric gene encoding all or a
CC portion of the cell cycle regulatory protein, in sense or antisense
CC orientation, where expression of the chimeric gene results in
CC production of altered levels of the cell cycle regulatory protein in
CC a transformed host cell. The nucleic acids and proteins may be
CC used to facilitate studies of cell cycle regulation in plants,
CC provide genetic tools to enhance cell growth in tissue culture,
CC increase gene transfer efficiency and provide more stable
CC transformations. The proteins may also provide targets to
CC facilitate design and/or identification of cell cycle regulatory
CC proteins that may be useful as herbicides.
XX
SQ Sequence 80 AA;

Query Match      57.9%; Score 44; DB 20; Length 80;
Best Local Similarity 43.8%; Pred. No. 0.54;
Matches 7; Conservative 6; Mismatches 3; Indels 0; Gaps 0;

QY 1 RYVDALNVLMMNNIIS 16
   |:|:|:|:|:|:|:|
Db 55 RIYDIVNVLESIGVLS 70

RESULT 23
AAW30502
ID AAW30502 standard; Peptide; 9 AA.
XX
AC AAW30502;
XX
DT 26-OCT-1998 (first entry)
XX
DE DP-1 transcription factor antagonist peptide motif.
XX
KW DP-1; transcription factor; antagonist; E2F protein; apoptosis;
KW cell proliferation; cardiovascular cell; restenosis; tumour;
KW surgical stent; therapy.
XX
OS Synthetic.
OS Homo sapiens.
XX
PN WO9828334-A1.
```

XX PD 02-JUL-1998.  
XX PF 22-DEC-1997; 97WO-CB03506.  
XX PR 20-DEC-1996; 96GB-0026589.  
XX PA (PROL-) PROLIFIX LTD.  
XX PI Bandara LR, La Thangue NB;  
XX PT WPI; 1998-377596/32.  
XX DR  
XX PT Polypeptide fragments of the DP-1 transcription factor - used for  
XX PT inducing apoptosis, specifically in tumour and cardiovascular cells,  
XX PT e.g. for preventing restenosis  
XX PS Claim 3; Page 44; 55pp; English.  
XX CC This peptide comprises amino acid residues 175-183 in the DEF box  
XX CC region (see AAW30501) of transcription factor DPL. Claimed peptides  
XX CC (II) (see AAW30502-07) containing this and/or another motif (see  
XX CC AAW30503) of the DEF box, are antagonists of the heterodimerisation  
XX CC of a DP protein with an E2F protein. Also claimed are variants  
XX CC of these peptides, especially containing substitutions of residues  
XX CC corresponding to residues 167, 169, 171 and 175 of DP-1, fusion  
XX CC proteins (III) comprising (I) or (II) and a membrane translocation  
XX CC sequence (see AAW30508), expression vectors encoding (I)-(III) and  
XX CC host cells. (I)-(III) are used therapeutically to induce apoptosis,  
XX CC specifically in tumour or cardiovascular cells, either in vivo or in  
XX CC vitro, e.g. for purging bone marrow. Surgical stents comprising  
XX CC (I)-(III) are used to treat or prevent restenosis in patients who  
XX CC have undergone angioplasty. (I)-(III) function by inactivating  
XX CC the DNA-binding activity of DP/E2F heterodimers. They are also  
XX CC used as research reagents, as positive controls in assays for  
XX CC identifying antagonists of DP-1/E2F dimerisation and as immunoassay  
XX CC agents. Also described is the use of sequences antisense to  
XX CC nucleic acids encoding (I)-(III) to control DP levels in cells,  
XX CC particularly by gene therapy. When formulated with cytotoxic  
XX CC or cytostatic agents, (I)-(III) enhance cell killing.  
XX SQ Sequence 9 AA;  
Query Match 55.3%; Score 42; DB 19; Length 9;  
Best Local Similarity 100.0%; Pred. No. 9.3e+05;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Qy 7 NVLMAMNII 15  
Db 1 NVLMAMNII 9  
RESULT 24  
AAU72601  
ID AAU72601 standard; Peptide; 35 AA.  
XX AC AAU72601;  
XX DT 26-FEB-2002 (first entry)  
XX DE DEF domain consensus sequence.  
XX KW Cell cycle protein; CCP; cell cycle regulation; herbicide;  
XX KW plant growth regulator; plant development; abiotic stress; biotic stress;  
XX KW nutrient deprivation; pathogen attack; crop yield; motif.  
XX OS Synthetic.  
XX PN WO200185946-A2.  
XX PD 15-NOV-2001.  
XX PF 14-MAY-2001; 2001WO-IB01307.

XX PR 12-MAY-2000; 2000US-204045P.  
XX PA (CROP-) CROPDESIGN NV.  
XX PI Inze D, Boudolf V, De Veylder L, Acosta JAT, Magyar Z;  
XX XX WPI; 2002-062249/08.  
XX DR  
XX PT New cell cycle protein and nucleic acid molecule encoding it useful for  
XX PT regulating cell cycle progression in plants and for identifying  
XX PT modulators which are useful as herbicides or plant growth regulators -  
XX PS Disclosure; Page 25; 316pp; English.  
XX CC The invention relates to a novel cell cycle protein (CCP) and the  
XX CC polynucleotides encoding them. CCP is useful for identifying a compound  
XX CC which modulates the activity of the polypeptide and which binds to the  
XX CC polypeptide and an anti-CCP antibody is useful for detecting the presence  
XX CC of CCP in a sample. A CCP modulator is useful for modulating the cell  
XX CC cycle or growth of a plant such as Arabidopsis thaliana, rice, wheat,  
XX CC maize, tomato, alfalfa, oilseed rape, soybean, sunflower and canola.  
XX CC CCP nucleic acid and polypeptide molecules are useful as modulating  
XX CC agents in regulating cell cycle progression in plants. CCP is useful to  
XX CC treat disorders characterised by insufficient or excessive production of  
XX CC CCP protein or production of CCP protein forms which have decreased or  
XX CC aberrant activity. Compounds that bind to or modulate the activity  
XX CC of CCP polypeptide are useful as herbicides or plant growth regulators.  
XX CC The polynucleotide is useful for modifying cell fate, plant development,  
XX CC plant morphology, biochemistry and/or physiology, the length of the G1,  
XX CC S, G2 and/or M phase of the cell cycle of a plant, initiation, promotion,  
XX CC stimulation or enhancement of cell division, DNA replication, seed set,  
XX CC seed size, seed development, tuber, fruit, leaf formation, shoot and root  
XX CC initiation and/or development, nodule function, dwarfism in plants,  
XX CC senescence, tolerance or resistance to stress. CCP, the polynucleotide  
XX CC and the anti-CCP antibody are useful in agriculture to modulate the  
XX CC protein levels or activity of a protein involved in the cell cycle due  
XX CC to environmental conditions, including abiotic stress such as  
XX CC cold, nutrient deprivation, heat, drought, salt stress, or biotic  
XX CC stress such as pathogen attack, to modulate e.g. enhance crop yields,  
XX CC and attenuate plant architecture, plant quality traits, plant  
XX CC reproduction and seed development, endoreduplication in storage cells,  
XX CC storage tissues and/or storage organs of plants or its parts. CCP is  
XX CC useful as an immunogen to generate antibodies. CCP protein is useful to  
XX CC screen for naturally occurring CCP substrates. The polynucleotide is  
XX CC useful for expressing CCP protein, to detect CCP mRNA, or a genetic  
XX CC lesion in a CCP gene and to modulate CCP activity. The present sequence  
XX CC represents a motif which may be found in a CCP protein of the invention.  
XX SQ Sequence 35 AA;  
Query Match 53.9%; Score 41; DB 23; Length 35;  
Best Local Similarity 66.7%; Pred. NO. 0.7;  
Matches 10; Conservative 0; Mismatches 5; Indels 0; Gaps 0;  
Qy 1 RYVDALNVLMAMNII 15  
Db 8 RVXDALNVLMAMNII 22  
RESULT 25  
AAW57054  
ID AAW57054 standard; peptide; 29 AA.  
XX AC AAW57054;  
XX DT 28-AUG-1998 (first entry)  
XX DE E2F activity inhibiting compound Ia-3.  
XX KW E2F activity; inhibitor; treatment; tumour; arteriosclerosis.  
XX OS Synthetic.



PT cell, for regulating proliferative capacity of a cell, involves  
PT modulating TERT transcription repression by Site C repressor binding  
PT site -  
PT  
PS Disclosure; Page 8; 66pp; English.  
XX  
CC The present invention relates to a new method of modulating expression  
CC of telomerase reverse transcriptase (TERT) from a TERT expression system  
CC that includes a TERT promoter and a Site C repressor binding site. The  
CC method of the invention involves modulating TERT transcription repression  
CC by the Site C repressor binding site. The method of the invention is  
CC useful for modulating expression of TERT for producing a mammalian  
CC antibody. The method is also useful in a variety of different  
CC applications, including immortalisation of cells, production of reagents  
CC for use in life science research, therapeutic applications, and  
CC therapeutic agent screening applications. Increasing TERT expression  
CC delays natural telomeric shortening and/or increases telomeric length and  
CC is useful for treating disease conditions such as Progeria or  
CC Hutchinson-Gilford syndrome, acquired immunodeficiency syndrome (AIDS),  
CC cardiovascular disease, osteoporosis, in skin rejuvenation and to inhibit  
CC immune senescence. The method can be employed to lengthen telomeres of  
CC osteoblast and osteoclast stem cells, encouraging bone replacement and  
CC proper remodeling and reinforcement, and can thus be used in bone marrow  
CC transplants for the treatment of cancer and skin grafts for burn  
CC victims and as such the method improves the survival and effectiveness of  
CC bone marrow and skin cell transplants. Decreasing TERT expression is  
CC useful for treating cellular proliferative disease conditions, including  
CC neoplastic disease conditions e.g. cancer. The present amino acid  
CC sequence represents the human DNA binding domain E2F-1.  
XX  
SQ Sequence 85 AA;  
  
Query Match 51.3%; Score 39; DB 23; Length 85;  
Best Local Similarity 43.8%; Pred. No. 5.7;  
Matches 7; Conservative 4; Mismatches 5; Indels 0; Gaps 0;  
  
QY 1 RYDNLNVLMMNNIIS 16  
Db 57 RYDITNVLEGIQLIA 72  
  
RESULT 28  
AAU78096  
ID AAU78096 standard; Protein; 76 AA.  
XX  
AC AAU78096;  
XX  
DT 18-JUN-2002 (first entry)  
XX  
DE Human DNA binding domain E2F-4.  
XX  
KW Human; telomerase reverse transcriptase; TERT; Site C; Progeria; burn;  
KW repressor binding site; Hutchinson-Gilford syndrome; AIDS; cancer;  
KW acquired immunodeficiency syndrome; cardiovascular disease; osteoporosis;  
KW skin rejuvenation; immune senescence; bone marrow transplant; skin graft;  
KW neoplastic disease; TERT minimal promoter; DNA binding domain; E2F-4.  
XX  
OS Homo sapiens.  
XX  
FH Key Location/Qualifiers  
FT Domain 7..49  
FT /label= DNA\_binding\_domain  
  
XX WO200216657-A1.  
XX  
PN 28-FEB-2002.  
XX  
PD 17-AUG-2001; 2001WO-US25861.  
XX  
PF 24-AUG-2000; 2000US-227865P.  
XX  
PR 01-SEP-2000; 2000US-230174P.  
XX  
PR 05-OCT-2000; 2000US-238345P.  
XX  
XX (SIERRA) SIERRA SCI INC.  
XX  
XX Andrews WH, Foster CA, Fraser S, Mohammadpour H;  
XX WPI; 2002-280952/32.  
XX  
XX Modulating expression of telomerase reverse transcriptase (TERT) in a  
XX cell, for regulating proliferative capacity of a cell, involves  
XX modulating TERT transcription repression by Site C repressor binding  
XX site -  
XX  
XX Disclosure; Page 8; 66pp; English.

PT cell, for regulating proliferative capacity of a cell, involves  
PT modulating TERT transcription repression by Site C repressor binding  
PT site -  
PT  
PS Disclosure; Page 8; 66pp; English.  
XX  
CC The present invention relates to a new method of modulating expression  
CC of telomerase reverse transcriptase (TERT) from a TERT expression system  
CC that includes a TERT promoter and a Site C repressor binding site. The  
CC method of the invention involves modulating TERT transcription repression  
CC by the Site C repressor binding site. The method of the invention is  
CC useful for modulating expression of TERT for producing a mammalian  
CC antibody. The method is also useful in a variety of different  
CC applications, including immortalisation of cells, production of reagents  
CC for use in life science research, therapeutic applications, and  
CC therapeutic agent screening applications. Increasing TERT expression  
CC delays natural telomeric shortening and/or increases telomeric length and  
CC is useful for treating disease conditions such as Progeria or  
CC Hutchinson-Gilford syndrome, acquired immunodeficiency syndrome (AIDS),  
CC cardiovascular disease, osteoporosis, in skin rejuvenation and to inhibit  
CC immune senescence. The method can be employed to lengthen telomeres of  
CC osteoblast and osteoclast stem cells, encouraging bone replacement and  
CC proper remodeling and reinforcement, and can thus be used in bone marrow  
CC transplants for the treatment of cancer and skin grafts for burn  
CC victims and as such the method improves the survival and effectiveness of  
CC bone marrow and skin cell transplants. Decreasing TERT expression is  
CC useful for treating cellular proliferative disease conditions, including  
CC neoplastic disease conditions e.g. cancer. The present amino acid  
CC sequence represents the human DNA binding domain E2F-1.  
XX  
SQ Sequence 85 AA;  
  
Query Match 51.3%; Score 39; DB 23; Length 85;  
Best Local Similarity 43.8%; Pred. No. 5.7;  
Matches 7; Conservative 4; Mismatches 5; Indels 0; Gaps 0;  
  
QY 1 RYDNLNVLMMNNIIS 16  
Db 57 RYDITNVLEGIQLIA 72  
  
RESULT 28  
ABB82986  
ID ABB82986 standard; Protein; 85 AA.  
XX  
AC ABB82986;  
XX  
DT 14-APR-2003 (first entry)  
XX  
DE Repressor protein E2F-1 DNA binding domain.  
XX  
KW Telomerase reverse transcriptase; TERT; Site C repressor; transcription;  
KW cytosolic; immunostimulant; anti-HIV; vulnery; telomerase; human;  
KW repressor protein; E2F-1.  
XX  
OS Homo sapiens.  
XX  
XX WO2002101010-A2.  
XX  
XX 19-DEC-2002.  
XX  
XX 06-JUN-2002; 2002WO-US17959.  
XX  
XX 07-JUN-2001; 2001US-296992P.  
XX  
XX (SIERRA) SIERRA SCI INC.  
XX  
XX Foster CA, Fraser S, Mohammadpour H, Andrews WH;  
XX WPI; 2003-167401/16.  
XX  
XX Modulating expression of telomerase reverse transcriptase TERT by  
XX blocking expression of TERT transcription, useful for the diagnosis and  
XX

CC The present invention relates to a new method of modulating expression  
 CC of telomerase reverse transcriptase (TERT) from a TERT expression system  
 CC that includes a TERT promoter and a Site C repressor binding site. The  
 CC method of the invention involves modulating TERT transcription repression  
 CC by the Site C repressor binding site. The method of the invention is  
 CC useful for modulating expression of TERT for producing a mammalian  
 CC antibody. The method is also useful in a variety of different  
 CC applications, including immortalisation of cells, production of reagents  
 CC for use in life science research, therapeutic applications, and  
 CC therapeutic agent screening applications. Increasing TERT expression  
 CC delays natural telomeric shortening and/or increases telomeric length and  
 CC is useful for treating disease conditions such as Progeria or  
 CC Hutchinson-Gilford syndrome, acquired immunodeficiency syndrome (AIDS),  
 CC cardiovascular disease, osteoporosis, in skin rejuvenation and to inhibit  
 CC immune senescence. The method can be employed to lengthen telomeres of  
 CC osteoblast and osteoclast stem cells, encouraging bone replacement and  
 CC proper remodeling and reinforcement, and can thus be used in bone marrow  
 CC transplants for the treatment of cancer and skin grafts for burn  
 CC victims and as such the method improves the survival and effectiveness of  
 CC bone marrow and skin cell transplants. Decreasing TERT expression is  
 CC useful for treating cellular proliferative disease conditions, including  
 CC neoplastic disease conditions e.g. cancer. The present amino acid  
 CC sequence represents the human DNA binding domain E2F-4.  
 XX  
 SQ Sequence 76 AA;

Query Match 50.0%; Score 38; DB 23; Length 76;  
 Best Local Similarity 46.7%; Pred. No. 7.7;  
 Matches 7; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

QY 1 RYVDALNVLMMNNII 15  
 | : | | | : : |  
 Db 47 RYVDITNVLEIGLI 61

## RESULT 30

ABB82987  
 ID ABB82987 standard; Protein; 76 AA.

XX ABB82987;

DT 14-APR-2003 (first entry)

XX Repressor protein E2F-4 DNA binding domain.

DE Telomerase reverse transcriptase; TERT; Site C repressor; transcription;  
 XX cytotostatic; immunostimulant; anti-HIV; vulneryary; telomerase; human;  
 KW repressor protein; E2F-4.

XX Homo sapiens.

XX WO2002101010-A2.

XX 19-DEC-2002.

XX 06-JUN-2002; 2002WO-US17959.

XX 07-JUN-2001; 2001US-296992P.

XX (SIER-) SIERRA SCI INC.

XX Foster CA, Fraser S, Mohammadpour H, Andrews WH;

XX WPI; 2003-167401/16.

XX Modulating expression of telomerase reverse transcriptase TERT by  
 PT blocking repression of TERT transcription, useful for the diagnosis and  
 PT treatment of disorders associated with aberrant telomerase activity  
 PT such as cancer and HIV -  
 XX

PS Disclosure; Page 7; 47pp; English.

XX The invention relates to modulating expression of telomerase reverse

CC transcriptase (TERT) expression system that includes a TERT promoter and  
 CC a Site C repressor binding site. The method involves modulating TERT  
 CC transcription repression by the Site C repressor binding site. The  
 CC methods and compositions of the present invention are useful for the  
 CC immortalization of cells, production of reagents in life science  
 CC research, therapeutic agent screening applications, diagnosis and  
 CC treatment of disorders associated with aberrant telomerase activity such  
 CC as cancer, progeria, immune senescence, HIV, and in skin rejuvenation.  
 CC The present sequence represents the DNA binding domain of a repressor  
 CC protein E2F-4.

XX Sequence 76 AA;

Query Match 50.0%; Score 38; DB 24; Length 76;  
 Best Local Similarity 46.7%; Pred. No. 7.7;  
 Matches 7; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

QY 1 RYVDALNVLMMNNII 15  
 | : | | | : : |  
 Db 47 RYVDITNVLEIGLI 61

## RESULT 31

AAR56646  
 ID AAR56646 standard; peptide; 24 AA.

XX AAR56646;

XX 25-MAR-2003 (updated)

DT 23-MAR-1995 (first entry)

XX Tyrosine activation motif mimic Fc-epsilon-RI-beta.

KW Tyrosine activation motif mimic; arthritis; asthma; allergy;  
 KW systemic lupus erythematosus; inflammatory bowe; disease;  
 KW allograft rejection; antinflammatory.

XX Synthetic.

XX WO9417095-A1.

XX 04-AUG-1994.

XX 28-JAN-1994; 94WO-US01025.

XX 29-JAN-1993; 93US-0013414.

XX (ARIA-) ARIAD PHARM INC.

XX Weiglele M, Tao G, Sundaramoorthi R, Dalgarno DC, Zydowsky LD;  
 PI Green J, Green OM;

XX WPI; 1994-264027/32.

XX New tyrosine activation motif analogues - for treating and  
 PT preventing immune and inflammatory disorders, e.g. allergies, by  
 PT inhibiting activation of mast cells, lymphocytes, etc.

XX Claim 10; Page 75; 94pp; English.

XX This TAM mimic may be used to treat or prevent conditions  
 CC associated with immune responses and inflammation, e.g. arthritis,  
 CC systemic lupus erythematosus, asthma, inflammatory bowel disease,  
 CC allergies, allograft rejection, etc. This peptide probably prevents  
 CC natural activation of target proteins so inhibit signal transduction.  
 CC (Updated on 25-MAR-2003 to correct PN field.)  
 CC (Updated on 25-MAR-2003 to correct PI field.)

XX Sequence 24 AA;

Query Match 48.7%; Score 37; DB 15; Length 24;  
 Best Local Similarity 63.6%; Pred. No. 2.6;  
 Matches 7; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

```
QY      1 RYVDALNVLMA 11
      |||: |||: |
Db      3 RYVEELNILSA 13

RESULT 32
AAY44493
ID      AAY44493 standard; peptide; 28 AA.
XX
XX      AAY44493;
XX
XX      27-MAR-2000 (first entry)
XX
XX      Wheat E2F derived peptide-1.
XX
KW      Wheat; transcription factor; TmE2F; plant retinoblastoma; plant E2F;
KW      Rb protein; cell cycle regulation; plant growth regulation;
KW      DNA replication; differentiation; E2F-derived peptide;
KW      DNA binding domain.
XX
XX      Triticum monococcum.
XX
XX      Key Location/Qualifiers
XX      Misc-difference 16 /label= Unknown
XX      Misc-difference 21 /label= Unknown
XX      Misc-difference 22 /label= Unknown
XX      Misc-difference 25 /label= Unknown
XX
XX      WO9958681-A2.
XX
XX      18-NOV-1999.
XX
XX      07-MAY-1999; 99WO-EP03159.
XX
XX      08-MAY-1998; 98ES-0000975.
XX      11-MAY-1998; 98ES-0000981.
XX
XX      (CNSJ ) CONSEJO SUPERIOR INVESTIGACIONES CIENTIF.
XX
XX      Gutierrez-Armenta C, Xie Q, Ramirez Parra E;
XX
XX      WPI; 2000-116290/10.
XX
XX      Regulating plant growth and/or cellular DNA replication and/or cell
XX      cycle progression, differentiation and development using a plant E2F
XX      polypeptide -
XX
XX      Claim 8; Page 26; 45pp; English.
XX
XX      The present sequence is a peptide comprising DNA-binding
XX      domain of wheat E2F protein that was identified
XX      in a wheat cDNA library by yeast two-hybrid screening. The E2F is a
XX      transcription factor that interacts with plant retinoblastoma (Rb) and
XX      is involved in cell cycle regulation. The present sequence
XX      has the ability to bind to E2F binding site in plant DNA and
XX      is used to control plant growth, cellular DNA replication, cell cycle
XX      progression, differentiation and development by altering its DNA-binding
XX      activity.
XX
XX      Sequence 28 AA;
XX
XX      Query Match 48.7%; Score 37; DB 21; Length 28;
XX      Best Local Similarity 46.7%; Pred. No. 3.2;
XX      Matches 7; Conservative 3; Mismatches 5; Indels 0; Gaps 0;
XX
QY      1 RYVDALNVLMA 15
      |||: |||: |
Db      4 RYDITNVLGIXL 18
```

```
RESULT 33
ABP00754
ID      ABP00754 standard; Protein; 88 AA.
XX
XX      ABP00754;
XX
XX      25-JUN-2002 (first entry)
XX
XX      Human ORFX protein sequence SEQ ID NO:1490.
XX
XX      Human; open reading frame; ORFX; gene therapy; cancer; cirrhosis;
XX      hyperproliferative disorder; psoriasis; benign tumour; haemorrhage;
XX      degenerative disorder; osteoarthritis; neurodegenerative disorder;
XX      cardiovascular disease; diabetes mellitus; systemic lupus erythematosus;
XX      hypertension; hypothyroidism; cholesterol ester storage disease;
XX      immune deficiency; immune disorder; infectious disease;
XX      autoimmune disorder; rheumatoid arthritis; autoimmune thyroiditis;
XX      myasthenia gravis.
XX
XX      Homo sapiens.
XX
XX      WO200192523-A2.
XX
XX      06-DEC-2001.
XX
XX      29-MAY-2001; 2001WO-US10836.
XX
XX      30-MAY-2000; 2000US-206132P.
XX      29-AUG-2000; 2000US-228716P.
XX
XX      (CURA-) CURAGEN CORP.
XX
XX      Shimkets RA, Leach MD;
XX
XX      WPI; 2002-106308/14.
XX      N-PSDB; ABN16506.
XX
XX      Novel human polypeptides and polynucleotides useful for diagnosing,
XX      preventing and treating cardiovascular disease, neurodegenerative,
XX      hyperproliferative disorders and autoimmune disorders -
XX
XX      Disclosure; SEQ ID 1490; 1037pp; English.
XX
XX      The present invention describes substantially purified human proteins
XX      (referred to as open reading frame, ORFX, where X is 1-11491 (see Table 1
XX      in the specification). ABN15762 to ABN27252 encode the human ORFX
XX      proteins given in ABP00010 to ABP11500. ORFX proteins are useful for
XX      treating or preventing a pathology associated with an ORFX-associated
XX      disorder in humans, and in the manufacture of a medicament for treating a
XX      syndrome associated with ORFX-associated disorder. ORFX polynucleotide
XX      sequences can be used in gene therapy. ORFX sequences can be used in the
XX      treatment of cancer, hyperproliferative disorders, cirrhosis of liver,
XX      osteoarthritis, neurodegenerative disorders, diabetes mellitus, systemic
XX      lupus erythematosus, hypertension, hypothyroidism, cholesterol ester
XX      storage disease, various immune deficiencies and disorders, infectious
XX      diseases, autoimmune disorders such as multiple sclerosis, rheumatoid
XX      arthritis, autoimmune thyroiditis, myasthenia gravis, graft-versus-host
XX      disease and autoimmune inflammatory eye disease. ORFX proteins are also
XX      useful for treating burns, incisions, ulcers, for treating osteoporosis,
XX      bone degenerative disorders, or periodontal disease, and for gut
XX      protection or regeneration and treatment of lung or liver fibrosis,
XX      reperfusion injury in various tissues and conditions resulting from
XX      systemic cytokine damage.
XX      N.B. The sequence data for this patent did not form part of the printed
XX      specification, but was obtained in electronic format directly from WIPO
XX      at ftp.wipo.int/pub/published_pct_sequences.
XX
XX      Sequence 88 AA;
```



```
PR 08-SEP-2000; 2000US-0232080.
PR 08-SEP-2000; 2000US-0232081.
PR 12-SEP-2000; 2000US-0231968.
PR 14-SEP-2000; 2000US-0232397.
PR 14-SEP-2000; 2000US-0232398.
PR 14-SEP-2000; 2000US-0232399.
PR 14-SEP-2000; 2000US-0232400.
PR 14-SEP-2000; 2000US-0232401.
PR 14-SEP-2000; 2000US-0233063.
PR 14-SEP-2000; 2000US-0233064.
PR 14-SEP-2000; 2000US-0233065.
PR 21-SEP-2000; 2000US-0234223.
PR 21-SEP-2000; 2000US-0234274.
PR 25-SEP-2000; 2000US-0234997.
PR 25-SEP-2000; 2000US-0234998.
PR 26-SEP-2000; 2000US-0235484.
PR 27-SEP-2000; 2000US-0235834.
PR 27-SEP-2000; 2000US-0235836.
PR 29-SEP-2000; 2000US-0236327.
PR 29-SEP-2000; 2000US-0236367.
PR 29-SEP-2000; 2000US-0236368.
PR 29-SEP-2000; 2000US-0236369.
PR 29-SEP-2000; 2000US-0236370.
PR 02-OCT-2000; 2000US-0236802.
PR 02-OCT-2000; 2000US-0237037.
PR 02-OCT-2000; 2000US-0237038.
PR 02-OCT-2000; 2000US-0237039.
PR 02-OCT-2000; 2000US-0237040.
PR 13-OCT-2000; 2000US-0239935.
PR 13-OCT-2000; 2000US-0239937.
PR 20-OCT-2000; 2000US-0240960.
PR 20-OCT-2000; 2000US-0241221.
PR 20-OCT-2000; 2000US-0241785.
PR 20-OCT-2000; 2000US-0241786.
PR 20-OCT-2000; 2000US-0241787.
PR 20-OCT-2000; 2000US-0241808.
PR 20-OCT-2000; 2000US-0241809.
PR 20-OCT-2000; 2000US-0241836.
PR 01-NOV-2000; 2000US-0244637.
PR 08-NOV-2000; 2000US-0246474.
PR 08-NOV-2000; 2000US-0246475.
PR 08-NOV-2000; 2000US-0246476.
PR 08-NOV-2000; 2000US-0246477.
PR 08-NOV-2000; 2000US-0246478.
PR 08-NOV-2000; 2000US-0246523.
PR 08-NOV-2000; 2000US-0246524.
PR 08-NOV-2000; 2000US-0246525.
PR 08-NOV-2000; 2000US-0246526.
PR 08-NOV-2000; 2000US-0246527.
PR 08-NOV-2000; 2000US-0246528.
PR 08-NOV-2000; 2000US-0246532.
PR 08-NOV-2000; 2000US-0246533.
PR 08-NOV-2000; 2000US-0246609.
PR 08-NOV-2000; 2000US-0246610.
PR 08-NOV-2000; 2000US-0246611.
PR 08-NOV-2000; 2000US-0246613.
PR 17-NOV-2000; 2000US-0249207.
PR 17-NOV-2000; 2000US-0249208.
PR 17-NOV-2000; 2000US-0249209.
PR 17-NOV-2000; 2000US-0249210.
PR 17-NOV-2000; 2000US-0249211.
PR 17-NOV-2000; 2000US-0249212.
PR 17-NOV-2000; 2000US-0249213.
PR 17-NOV-2000; 2000US-0249214.
PR 17-NOV-2000; 2000US-0249215.
PR 17-NOV-2000; 2000US-0249216.
PR 17-NOV-2000; 2000US-0249217.
PR 17-NOV-2000; 2000US-0249218.
PR 17-NOV-2000; 2000US-0249244.
PR 17-NOV-2000; 2000US-0249245.
PR 17-NOV-2000; 2000US-0249264.
PR 17-NOV-2000; 2000US-0249265.
PR 17-NOV-2000; 2000US-0249297.
PR 17-NOV-2000; 2000US-0249299.
PR 17-NOV-2000; 2000US-0249299.

PR 17-NOV-2000; 2000US-0249300.
PR 01-DEC-2000; 2000US-0250160.
PR 01-DEC-2000; 2000US-0250391.
PR 05-DEC-2000; 2000US-0251030.
PR 05-DEC-2000; 2000US-0251988.
PR 05-DEC-2000; 2000US-0256719.
PR 06-DEC-2000; 2000US-0251479.
PR 08-DEC-2000; 2000US-0251856.
PR 08-DEC-2000; 2000US-0251868.
PR 08-DEC-2000; 2000US-0251869.
PR 08-DEC-2000; 2000US-0251989.
PR 08-DEC-2000; 2000US-0251990.
PR 11-DEC-2000; 2000US-0254097.
PR 05-JAN-2001; 2001US-0259678.
XX (HUMA-) HUMAN GENOME SCI INC.
PA Rosen CA, Barash SC, Ruben SM;
PI WPI; 2001-465567/50.
DR N-PSDB; AAS39447.
DR XX
DR XX
XX Isolated polypeptide for treating, preventing and/ or prognosing
PT disorders related to the colon including colon cancers and also for
PT testing and detection e.g. diagnosis -
XX
PS Claim 11; SEQ ID No 344; 562pp; English.
XX
CC The present invention relates to the isolation of novel human colon
CC associated polypeptides, and the cDNA (AAS39348-AAS39581) and genomic
CC sequences encoding for them. The sequences of the invention are useful
CC in the diagnosis, treatment, prevention and/or prognosis of disorders
CC of the colon including colon cancer, congenital abnormalities
CC (e.g. atresia and stenosis), bacterial and viral infections,
CC inflammatory bowel disease (IBD), neoplastic cell disorders
CC (e.g. polyps and adenomas, intestinal inflammatory disorders, colitis,
CC colonic inflammation, diarrhoea and dysentery, malabsorption syndromes
CC (e.g. lactose intolerance), intestinal obstruction and sigmoid diseases.
CC The polynucleotide sequences of the invention can also be used in gene
CC therapy. AAU22468-AAU22701 represent the novel human colon associated
CC polypeptides of the invention.
CC Note: The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ Sequence 63 AA;
Query Match 46.1%; Score 35; DB 22; Length 63;
Best Local Similarity 57.1%; Pred. No. 23;
Matches 8; Conservative 1; Mismatches 5; Indels 0; Gaps 0;
QY 2 VYDALNVLMMNII 15
DB 8 VYFLNFLMLKNLI 21
|| || || || ||
RESULT 36
AAM92530
ID AAM92530 standard; Protein; 63 AA.
XX
XX AAM92530;
XX
XX 06-NOV-2001 (first entry)
DE Human digestive system antigen SEQ ID NO: 1879.
XX
XX Human; digestive system antigen; gene therapy; cancer; appendicitis;
KW ulcerative colitis; infection; Hirschsprung's disease; chronic colitis;
KW digestive system disorder; Meckel's diverticulum.
XX
XX Homo sapiens.
XX
XX WO200155314-A2.
XX
XX
```



XX	02-AUG-2001.	29-SEP-2000; 2000US-0236367.	PR
PD		29-SEP-2000; 2000US-0236368.	PR
XX		29-SEP-2000; 2000US-0236369.	PR
XX		29-SEP-2000; 2000US-0236370.	PR
PF	17-JAN-2001; 2001WO-US01324.	29-SEP-2000; 2000US-0236370.	PR
XX		02-OCT-2000; 2000US-0236802.	PR
XX		02-OCT-2000; 2000US-0237037.	PR
PR		02-OCT-2000; 2000US-0237038.	PR
PR		02-OCT-2000; 2000US-0237039.	PR
PR		02-OCT-2000; 2000US-0237040.	PR
PR		13-OCT-2000; 2000US-0239935.	PR
PR		13-OCT-2000; 2000US-0239937.	PR
PR		13-OCT-2000; 2000US-0240960.	PR
PR		20-OCT-2000; 2000US-0241221.	PR
PR		20-OCT-2000; 2000US-0241785.	PR
PR		20-OCT-2000; 2000US-0241786.	PR
PR		20-OCT-2000; 2000US-0241787.	PR
PR		20-OCT-2000; 2000US-0241808.	PR
PR		20-OCT-2000; 2000US-0241809.	PR
PR		20-OCT-2000; 2000US-0241826.	PR
PR		01-NOV-2000; 2000US-0244617.	PR
PR		08-NOV-2000; 2000US-0246474.	PR
PR		08-NOV-2000; 2000US-0246475.	PR
PR		08-NOV-2000; 2000US-0246476.	PR
PR		08-NOV-2000; 2000US-0246477.	PR
PR		08-NOV-2000; 2000US-0246478.	PR
PR		08-NOV-2000; 2000US-0246523.	PR
PR		08-NOV-2000; 2000US-0246524.	PR
PR		08-NOV-2000; 2000US-0246525.	PR
PR		08-NOV-2000; 2000US-0246526.	PR
PR		08-NOV-2000; 2000US-0246527.	PR
PR		08-NOV-2000; 2000US-0246528.	PR
PR		08-NOV-2000; 2000US-0246532.	PR
PR		08-NOV-2000; 2000US-0246610.	PR
PR		08-NOV-2000; 2000US-0246611.	PR
PR		08-NOV-2000; 2000US-0246613.	PR
PR		17-NOV-2000; 2000US-0249207.	PR
PR		17-NOV-2000; 2000US-0249208.	PR
PR		17-NOV-2000; 2000US-0249209.	PR
PR		17-NOV-2000; 2000US-0249210.	PR
PR		17-NOV-2000; 2000US-0249211.	PR
PR		17-NOV-2000; 2000US-0249212.	PR
PR		17-NOV-2000; 2000US-0249213.	PR
PR		17-NOV-2000; 2000US-0249214.	PR
PR		17-NOV-2000; 2000US-0249215.	PR
PR		17-NOV-2000; 2000US-0249216.	PR
PR		17-NOV-2000; 2000US-0249217.	PR
PR		17-NOV-2000; 2000US-0249218.	PR
PR		17-NOV-2000; 2000US-0249244.	PR
PR		17-NOV-2000; 2000US-0249245.	PR
PR		17-NOV-2000; 2000US-0249264.	PR
PR		17-NOV-2000; 2000US-0249265.	PR
PR		17-NOV-2000; 2000US-0249297.	PR
PR		17-NOV-2000; 2000US-0249299.	PR
PR		17-NOV-2000; 2000US-0249300.	PR
PR		01-DEC-2000; 2000US-0250160.	PR
PR		01-DEC-2000; 2000US-0250391.	PR
PR		05-DEC-2000; 2000US-0251030.	PR
PR		05-DEC-2000; 2000US-0251988.	PR
PR		05-DEC-2000; 2000US-0256719.	PR
PR		06-DEC-2000; 2000US-0251479.	PR
PR		08-DEC-2000; 2000US-0251856.	PR
PR		08-DEC-2000; 2000US-0251868.	PR
PR		08-DEC-2000; 2000US-0251869.	PR
PR		08-DEC-2000; 2000US-0251989.	PR
PR		11-DEC-2000; 2000US-0251990.	PR
PR		11-DEC-2000; 2000US-0254097.	PR
PR		05-JAN-2001; 2001US-0259678.	XX
XX		(HUMA-) HUMAN GENOME SCI INC.	PA
XX		Rosen CA, Barash SC, Ruben SM;	XX
PI			PI
XX			XX



PR 01-SEP-2000; 2000US-0229344.  
PR 01-SEP-2000; 2000US-0229345.  
PR 05-SEP-2000; 2000US-0229509.  
PR 05-SEP-2000; 2000US-0229513.  
PR 06-SEP-2000; 2000US-0230437.  
PR 06-SEP-2000; 2000US-0230438.  
PR 08-SEP-2000; 2000US-0231242.  
PR 08-SEP-2000; 2000US-0231243.  
PR 08-SEP-2000; 2000US-0231244.  
PR 08-SEP-2000; 2000US-0231413.  
PR 08-SEP-2000; 2000US-0231414.  
PR 08-SEP-2000; 2000US-0232080.  
PR 08-SEP-2000; 2000US-0232081.  
PR 12-SEP-2000; 2000US-0231968.  
PR 14-SEP-2000; 2000US-0232397.  
PR 14-SEP-2000; 2000US-0232398.  
PR 14-SEP-2000; 2000US-0233064.  
PR 14-SEP-2000; 2000US-0233065.  
PR 21-SEP-2000; 2000US-0234223.  
PR 21-SEP-2000; 2000US-0234274.  
PR 25-SEP-2000; 2000US-0234597.  
PR 25-SEP-2000; 2000US-0234998.  
PR 26-SEP-2000; 2000US-0235484.  
PR 27-SEP-2000; 2000US-0235834.  
PR 27-SEP-2000; 2000US-0235836.  
PR 28-SEP-2000; 2000US-0236327.  
PR 28-SEP-2000; 2000US-0236367.  
PR 29-SEP-2000; 2000US-0236368.  
PR 29-SEP-2000; 2000US-0236369.  
PR 29-SEP-2000; 2000US-0236370.  
PR 02-OCT-2000; 2000US-0236802.  
PR 02-OCT-2000; 2000US-0237037.  
PR 02-OCT-2000; 2000US-0237038.  
PR 02-OCT-2000; 2000US-0237039.  
PR 02-OCT-2000; 2000US-0237040.  
PR 13-OCT-2000; 2000US-0239935.  
PR 13-OCT-2000; 2000US-0239937.  
PR 20-OCT-2000; 2000US-0240960.  
PR 20-OCT-2000; 2000US-0241221.  
PR 20-OCT-2000; 2000US-0241785.  
PR 20-OCT-2000; 2000US-0241786.  
PR 20-OCT-2000; 2000US-0241787.  
PR 20-OCT-2000; 2000US-0241808.  
PR 20-OCT-2000; 2000US-0241809.  
PR 20-OCT-2000; 2000US-0241826.  
PR 01-NOV-2000; 2000US-0244617.  
PR 08-NOV-2000; 2000US-0246474.  
PR 08-NOV-2000; 2000US-0246475.  
PR 08-NOV-2000; 2000US-0246476.  
PR 08-NOV-2000; 2000US-0246477.  
PR 08-NOV-2000; 2000US-0246478.  
PR 08-NOV-2000; 2000US-0246523.  
PR 08-NOV-2000; 2000US-0246524.  
PR 08-NOV-2000; 2000US-0246525.  
PR 08-NOV-2000; 2000US-0246526.  
PR 08-NOV-2000; 2000US-0246527.  
PR 08-NOV-2000; 2000US-0246528.  
PR 08-NOV-2000; 2000US-0246532.  
PR 08-NOV-2000; 2000US-0246609.  
PR 08-NOV-2000; 2000US-0246610.  
PR 08-NOV-2000; 2000US-0246611.  
PR 08-NOV-2000; 2000US-0246613.  
PR 17-NOV-2000; 2000US-0249207.  
PR 17-NOV-2000; 2000US-0249208.  
PR 17-NOV-2000; 2000US-0249209.  
PR 17-NOV-2000; 2000US-0249210.  
PR 17-NOV-2000; 2000US-0249211.  
PR 17-NOV-2000; 2000US-0249212.  
PR 17-NOV-2000; 2000US-0249213.

PR 17-NOV-2000; 2000US-0249214.  
PR 17-NOV-2000; 2000US-0249215.  
PR 17-NOV-2000; 2000US-0249216.  
PR 17-NOV-2000; 2000US-0249217.  
PR 17-NOV-2000; 2000US-0249218.  
PR 17-NOV-2000; 2000US-0249244.  
PR 17-NOV-2000; 2000US-0249245.  
PR 17-NOV-2000; 2000US-0249264.  
PR 17-NOV-2000; 2000US-0249265.  
PR 17-NOV-2000; 2000US-0249297.  
PR 17-NOV-2000; 2000US-0249299.  
PR 17-NOV-2000; 2000US-0249300.  
PR 01-DEC-2000; 2000US-0250160.  
PR 01-DEC-2000; 2000US-0250391.  
PR 05-DEC-2000; 2000US-0251030.  
PR 05-DEC-2000; 2000US-0251988.  
PR 06-DEC-2000; 2000US-0256719.  
PR 06-DEC-2000; 2000US-0251479.  
PR 08-DEC-2000; 2000US-0251856.  
PR 08-DEC-2000; 2000US-0251869.  
PR 08-DEC-2000; 2000US-0251989.  
PR 08-DEC-2000; 2000US-0251990.  
PR 11-DEC-2000; 2000US-0254097.  
PR 03-JAN-2001; 2001US-0259678.

XX (HUMA-) HUMAN GENOME SCI INC.

XX Rosen CA, Barash SC, Ruben SM;

XX WPI; 2001-483426/52.

DR N-PSDB; AAKG2143.

XX Nucleic acids encoding human immune/hematopoietic antigen polypeptides,  
PT useful for preventing, diagnosing and/or treating cancers and  
PT metastasis -

PS Claim 11; SEQ ID NO 16955; 3071pp + Sequence Listing; English.

XX AAK54951 to AAK64702 encode the human immune/hematopoietic antigen (I)  
CC amino acid sequences given in AAK82170 to AAK91921. (I) have cytostatic  
CC activity, and can be used in gene therapy and vaccine production. (I)  
CC proteins and polynucleotides may be used in the prevention, diagnosis and  
CC treatment of diseases associated with inappropriate (I) expression. For  
CC example, they may be used to treat disorders associated with decreased  
CC expression by rectifying mutations or deletions in a patient's genome  
CC that affect the activity of (I) by expressing inactive proteins or to  
CC supplement the patient's own production of (I). Additionally, (I)  
CC polynucleotides may be used to produce the secreted (I), by inserting the  
CC the nucleic acids into a host cell and culturing the cell to express the  
CC protein. (I) proteins and polynucleotides may be used to prevent,  
CC diagnose and treat immune/hematopoietic-related diseases, especially  
CC cancers and cancer metastases of hematopoietic-derived cells. AAK64703  
CC to AAK87694 represent human immune/hematopoietic antigen genomic  
CC sequences from the present invention. AAK54942 to AAK54950 and AAK82169  
CC represent sequences used in the exemplification of the present invention.

XX Sequence 87 AA;

Query Match 46.1%; Score 35; DB 22; Length 87;

Best Local Similarity 58.3%; Pred. No. 36;

Matches 7; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY 3 YDALNVLMMNII 14

Db 73 YDILNVLMMNII 84

RESULT 39

AA53284

ID AAB53284 standard; Protein; 90 AA.

XX

AC AAB53284;

XX 09-MAR-2001 (first entry)  
KW Human colon cancer antigen protein sequence SEQ ID NO:824.  
DE  
XX  
KW Human; colon cancer; colon cancer antigen; diagnosis; detection;  
KW identification; cytostatic; cardioactive; neuroprotective; vulnary;  
KW immunomodulatory; muscular; gynaecological; gastrointestinal;  
KW nephrotropic; antiinfective; antibacterial; gene therapy; wound;  
KW neural disorder; immune system disorder; muscular disorder;  
KW reproductive disorder; gastrointestinal disorder; renal disorder;  
KW infectious disease; cardiovascular disorder.  
XX  
OS Homo sapiens.  
XX  
XX WO200055351-A1.  
XX  
XX 21-SEP-2000.  
XX  
XX 08-MAR-2000; 2000WO-US05883.  
XX  
XX 12-MAR-1999; 99US-0124270.  
XX (HUMA-) HUMAN GENOME SCI INC.  
XX  
XX Rosen CA, Ruben SM;  
XX  
XX WPI; 2000-587534/55.  
XX N-PSDB; AAC98041.  
XX  
XX Colon cancer associated gene sequences, referred to as colon cancer  
XX antigens, useful for the treatment, prevention, and diagnosis of colon  
XX disorders such as colon cancer -  
XX  
XX Claim 11; Page 1376; 2104pp; English.  
XX  
XX AAC97991 to AAC98763 encode the human colon cancer associated proteins,  
XX called human colon cancer antigens, given in AAB51234 to AAB54006. The  
XX human colon cancer antigens can have cytostatic, cardioactive, muscular;  
XX neuroprotective, immunomodulatory, gynaecological, gastrointestinal, and  
XX vulnary, nephrotropic, antiinfective and antibacterial activities, and  
XX can be used in gene therapy. The colon cancer antigen polynucleotides,  
XX proteins and antibodies to the proteins are useful for the prevention,  
XX treatment and diagnosis of colon disorders, such as colon cancer. The  
XX polynucleotides may be used in diagnostics and research, such as for  
XX chromosome identification, and as hybridisation probes. The proteins  
XX may also be used to prevent diseases such as neural disorders, immune  
XX system disorders, muscular disorders, reproductive disorders,  
XX gastrointestinal disorders, wounds, renal disorders, infectious  
XX diseases, and cardiovascular disorders. AAC98764 to AAC98772 and  
XX AAB54007 represent sequences used in the exemplification of the present  
XX invention.  
XX  
SQ Sequence 90 AA;  
Query Match 46.1%; Score 35; DB 21; Length 90;  
Best Local Similarity 75.0%; Pred. No. 38;  
Matches 6; Conservative 2; Mismatches 0; Indels 0; Gaps 0;  
QY 3 YDALNVLM 10  
:|||||:  
DB 42 FDALNLM 49  
RESULT 40  
AAW79339  
ID AAW79339 standard; Protein; 44 AA.  
XX  
XX AAW79339;  
AC  
XX 24-NOV-1998 (first entry)  
DE Staphylococcus aureus haem lipoic acid synthetase.

XX Haemophilus influenzae; haem lipoic acid synthetase;  
KW lipoate biosynthesis protein A; lipA; treatment; prevention;  
KW bacterial infection; Helicobacter pylori; vaccine.  
XX  
OS Staphylococcus aureus.  
XX  
XX WO9823738-A2.  
XX  
XX 04-JUN-1998.  
XX  
XX 24-NOV-1997; 97WO-US22092.  
XX  
XX 25-NOV-1996; 96US-0031469.  
XX (SMIK ) SMITHKLINE BEECHAM CORP.  
XX  
XX Warren RL;  
XX  
XX WPI; 1998-322718/28.  
XX N-PSDB; AAV59884.  
XX  
XX New nucleic acid from Staphylococcus aureus NCIMB 40771 - useful  
XX for, e.g. diagnosis, prevention and treatment of bacterial  
XX infection(s)  
XX  
XX Claim 5; Pages 51; 114pp; English.  
XX  
XX AAW71314 and AAW79337-39 represent Staphylococcus aureus WCHU (NCIMB  
XX 40771) proteins that have homology to Haemophilus influenzae haem  
XX lipoic acid synthetase (lipoate biosynthesis protein A (lipA) homologue.  
XX The protein is used to generate antibodies and to screen for  
XX antimicrobials. The products are used to treat or prevent bacterial  
XX infections, particularly where caused by S. aureus but also  
XX against Helicobacter pylori. Particular applications are to treat  
XX subjects before surgery or insertion of an in-dwelling device  
XX (alternatively the device itself is impregnated before placement). The  
XX nucleic acid sequence is used as sources of antisense sequences (for  
XX therapeutic use) or regulatory elements for controlling expression of  
XX bacterial genes, and for antibacterial screening. The protein can be  
XX also used as a vaccine.  
XX  
SQ Sequence 44 AA;  
Query Match 44.7%; Score 34; DB 19; Length 44;  
Best Local Similarity 77.8%; Pred. No. 23;  
Matches 7; Conservative 0; Mismatches 2; Indels 0; Gaps 0;  
QY 3 YDALNVLM 11  
:|||||:  
DB 10 YDALETLM 18  
RESULT 41  
AAW65070  
ID AAW65070 standard; Protein; 51 AA.  
XX  
XX AAW65070;  
AC  
XX 01-FEB-2000 (first entry)  
DT  
XX Human 5' EST related polypeptide SEQ ID NO:1231.  
DE  
XX Human; 5' EST; expressed sequence tag; secreted protein; diagnosis;  
KW gene therapy; chromosome mapping; upstream regulatory sequence;  
KW forensic; location; development; protein synthesis; stability;  
KW regulation; identification.  
XX  
XX Homo sapiens.  
XX  
XX WO9953051-A2.  
XX  
XX 21-OCT-1999.  
PD

XX 09-APR-1999; 99WO-IB00712.  
 XX 09-APR-1998; 98US-0057719.  
 PR 28-APR-1998; 98US-0069047.  
 XX (GEST ) GENSET.  
 PA Dumas Milne Edwards J, Duclert A, Giordano J;  
 PI WPI; 2000-038446/03.  
 XX DR N-PSDB; AA242684.  
 XX Novel secreted protein 5' expressed sequence tag sequences used in  
 PT diagnostic, forensic, gene therapy, and chromosome mapping procedures  
 PT Claim 3; Page 724; 837pp; English.  
 XX AA242265 to AA243075 represent novel 5' expressed sequence tag (EST)  
 CC sequences, corresponding to human secreted proteins. AA242651 to  
 CC AA242653 represent the EST-related proteins corresponding to AA242265 to  
 CC AA243052. The 5' ESTs can be used for producing secreted human gene  
 CC products. They can be used to identify and isolate 5' untranslated  
 CC regions (UTRs) and upstream regulatory regions which control the  
 CC location, development stage, rate, and quantity of protein synthesis, as  
 CC well as stability of mRNA. The ESTs are also useful as probes for  
 CC chromosome mapping, and to obtain full length cDNA clones. The ESTs can  
 CC also be used in forensic procedures to identify individuals, or in  
 CC diagnostic procedures to identify individuals having genetic diseases  
 CC resulting from abnormal gene expression. The products may also be used in  
 CC gene therapy protocols. The nucleic acids encoding signal peptides can be  
 CC used for directing extracellular secretion of a polypeptide or the  
 CC insertion of a polypeptide into a membrane, or importing a polypeptide  
 CC into a cell. The proteins encoded by the EST sequences may be useful in  
 CC treating a variety of human conditions. Secreted proteins have  
 CC therapeutic value, and the identification of new secreted proteins is  
 CC valuable. AA242249 to AA242264 and AA242644 to AA242650 represent  
 CC sequences used in the exemplification of the present invention.  
 XX SQ Sequence 51 AA;  
 SQ Query Match 44.7%; Score 34; DB 21; Length 51;  
 Best Local Similarity 42.9%; Pred. No. 28;  
 Matches 6; Conservative 2; Mismatches 6; Indels 0; Gaps 0;  
 QY 3 YDALNVLMMNIIIS 16  
 DB 36 YDPINWYNNVFS 49  
 RESULT 42  
 AAB80361  
 ID AAB80361 standard; Protein; 59 AA.  
 AC AAB80361;  
 XX 24-APR-2001 (first entry)  
 DT Human prostate cancer antigen #89.  
 XX Immunosuppressive; neutrotropic; neuroprotective; antiviral; vulnary;  
 KW anticonvulsant; antibacterial; antifungal; antiparasitic; cardiac;  
 KW immune disorder; cardiovascular disorder; neurological disease;  
 KW infection; cancer; cytostatic; antiarthritic; antirheumatic;  
 KW antiasthmatic; anticonvulsant; vasotropic; vulnary; human;  
 KW secreted protein; prostate cancer antigen.  
 XX Homo sapiens.  
 OS WO200107476-A1.  
 PN 01-FEB-2001.  
 XX PD

PF 20-JUL-2000; 2000WO-US19666.  
 XX 21-JUL-1999; 99US-0144972.  
 PR 13-AUG-1999; 99US-0148681.  
 PR 17-AUG-1999; 99US-0149173.  
 PR 06-OCT-1999; 99US-0158004.  
 PR 05-APR-2000; 2000US-0194689.  
 XX (HUMA-) HUMAN GENOME SCI INC.  
 PA (ROSE/) ROSEN C A.  
 XX Rosen CA, Birse C;  
 PI WPI; 2001-138554/14.  
 XX New nucleic acid molecule encoding human secreted prostate cancer  
 PT antigens, useful for the diagnosis and treatment of disorders such as  
 PT cancer, leukemia and autoimmune disease -  
 XX Disclosure; Page 56; 433pp; English.  
 XX The present invention relates to human secreted prostate cancer antigen  
 CC coding sequences (AAF72741-AAF72789) and proteins (AAB80273-AAB80321). for  
 CC The coding sequences and proteins of the present invention are useful for  
 CC preventing, treating or ameliorating a medical condition; and for the  
 CC diagnosis and treatment of diseases and disorders. Diseases and disorders  
 CC that can be diagnosed and treated include (auto)immune diseases (e.g.  
 CC graft versus host disease and rheumatoid arthritis), inflammatory and  
 CC allergic disorders (e.g. asthma, hyperproliferative disorders (e.g.  
 CC cancers and leukemias), cardiovascular disorders (e.g. heart attacks and  
 CC arrhythmias), cerebrovascular disorders (e.g. stroke), arterial occlusive  
 CC disorders (e.g. arteriosclerosis), angiogenesis related disorders (e.g.  
 CC retinopathy and keloid scars), ocular disorders (e.g. glaucoma),  
 CC neurological disorders (e.g. Alzheimer's, Parkinson's disease, epilepsy  
 CC and Creutzfeld-Jacob disease) and infections caused by bacteria, fungi,  
 CC viruses or parasites. They may also be useful for wound healing,  
 CC epithelial cell proliferation, supporting cell culture, tissue  
 CC regeneration, birth control and as a food additive or preservative.  
 XX SQ Sequence 59 AA;  
 SQ Query Match 44.7%; Score 34; DB 22; Length 59;  
 Best Local Similarity 50.0%; Pred. No. 34;  
 Matches 6; Conservative 5; Mismatches 1; Indels 0; Gaps 0;  
 QY 1 RYVDALNVLMMAM 12  
 DB 38 RYVEYLNLLISV 49  
 RESULT 43  
 AAU60610  
 ID AAU60610 standard; Protein; 60 AA.  
 XX AAU60610;  
 AC 27-FEB-2002 (first entry)  
 DT Propionibacterium acnes immunogenic protein #21506.  
 XX SAPHO syndrome; synovitis; acne; pustulosis; hypertosis; osteomyelitis;  
 KW uveitis; endophthalmitis; bone; joint; central nervous system; ELISA;  
 KW inflammatory lesion; acne vulgaris; enzyme linked immunosorbent assay;  
 KW dermatological; osteopathic; neuroprotectant.  
 XX Propionibacterium acnes.  
 OS WO200181581-A2.  
 PN 01-NOV-2001.  
 XX PD 20-APR-2001; 2001WO-US12865.  
 XX PF

PR 21-APR-2000; 2000US-199047P.  
PR 02-JUN-2000; 2000US-208841P.  
PR 07-JUL-2000; 2000US-216747P.  
XX  
XX  
PA (CORI-) CORIXA CORP.  
XX  
XX Skeiky YAM, Persing DH, Mitcham JL, Wang SS, Bhatia A;  
PI L'maisonneuve J, Zhang Y, Jen S, Carter D;  
XX  
XX WPI; 2001-616774/71.  
DR N-PSDB; AAS59610.  
XX  
XX Propionibacterium acnes polypeptides and nucleic acids useful for  
PT vaccinating against and diagnosing infections, especially useful for  
PT treating acne vulgaris -  
XX  
XX Example 1; SEQ ID No 21805; 1069pp; English.  
XX  
XX Sequences AAU39105-AAUG8017 represent Propionibacterium acnes immunogenic  
CC polypeptides. The proteins and their associated DNA sequences are used in  
CC the treatment, prevention and diagnosis of medical conditions caused by  
CC P. acnes. The disorders include SAPHO syndrome (synovitis, acne,  
CC pustulosis, hypertosis and osteomyelitis), uveitis and endophthalmitis.  
CC P. acnes is also involved in infections of bone, joints and the central  
CC nervous system, however it is particularly involved in the inflammatory  
CC lesions associated with acne vulgaris. A method for detecting the  
CC presence or absence of P. acnes in a patient comprises contacting a  
CC sample with a binding agent that binds to the proteins of the invention  
CC and determining the amount of bound protein in the sample. The  
CC polypeptides may be used as antigens in the production of antibodies  
CC specific for P. acnes proteins. These antibodies can be used to  
CC downregulate expression and activity of P. acnes polypeptides and  
CC therefore treat P. acnes infections. The antibodies may also be used as  
CC diagnostic agents for determining P. acnes presence, for example, by  
CC enzyme linked immunosorbent assay (ELISA).  
CC Note: The sequence data for this patent did not form part of the printed  
CC specification, but was obtained in electronic format directly from WIPO  
CC at ftp.wipo.int/pub/published\_pct\_sequences.  
XX  
XX SQ Sequence 60 AA;  
Query Match 44.7%; Score 34; DB 22; Length 60;  
Best Local Similarity 46.7%; Pred. No. 35;  
Matches 7; Conservative 4; Mismatches 4; Indels 0; Gaps 0;  
Qy 2 VYDALNVLMMNNIIS 16  
Db 24 VYDSLNMGLRIHS 38  
|||:|:|:|:|  
|||:|:|:|:|  
RESULT 44  
AAM58362  
ID AAM58362 standard; Protein; 67 AA.  
XX  
XX AAM58362;  
AC  
DT 05-NOV-2001 (first entry)  
XX  
XX Human brain expressed single exon probe encoded protein SEQ ID NO: 30467.  
DE  
XX Human; brain expressed exon; gene expression analysis; probe;  
KW microarray; Alzheimer's disease; multiple sclerosis; schizophrenia;  
KW epilepsy; cancer.  
XX  
XX Homo sapiens.  
OS  
XX WO200157275-A2.  
PN  
XX  
PD 09-AUG-2001.  
XX  
XX 30-JAN-2001; 2001WO-US00667.  
PF  
XX 04-FEB-2000; 2000US-0180312.  
PR

PR 26-MAY-2000; 2000US-0207456.  
PR 30-JUN-2000; 2000US-0608408.  
PR 03-AUG-2000; 2000US-0632366.  
PR 21-SEP-2000; 2000US-0234687.  
PR 27-SEP-2000; 2000US-0236359.  
PR 04-OCT-2000; 2000GB-0024263.  
XX  
XX (MOLE-) MOLECULAR DYNAMICS INC.  
XX  
XX Penn SG, Hanzel DK, Chen W, Rank DR;  
PI  
XX WPI; 2001-483446/52.  
DR  
XX Single exon nucleic acid probes for analyzing gene expression in human  
PT brains -  
PT  
XX Example 4; SEQ ID NO: 30467; 650pp + Sequence Listing; English.  
PS  
XX The present invention provides a number of single exon nucleic acid  
CC probes which are derived from genomic sequences expressed in the human  
CC brain. They can be used to measure gene expression in brain cell samples,  
CC which may enable the diagnosis and improved treatment of nervous system  
CC diseases such as Alzheimer's disease, multiple sclerosis, schizophrenia,  
CC epilepsy and cancers. The present sequence is a protein encoded by one of  
CC the probes of the invention.  
XX  
XX SQ Sequence 67 AA;  
Query Match 44.7%; Score 34; DB 22; Length 67;  
Best Local Similarity 35.7%; Pred. No. 40;  
Matches 5; Conservative 6; Mismatches 3; Indels 0; Gaps 0;  
Qy 2 VYDALNVLMMNNI 15  
Db 38 IITTNIIAIINII 51  
|||:|:|:|:|  
|||:|:|:|:|  
RESULT 45  
AAU07882  
ID AAU07882 standard; Protein; 80 AA.  
XX  
XX AAU07882;  
AC  
DT 19-DEC-2001 (first entry)  
XX  
XX Polypeptide sequence for mammalian Sp987.  
DE  
XX Mammalian; reproductive-specific protein; male infertility;  
KW spermatogenesis; sperm count disorder; anti infertility; reproduction.  
KW Mammalia.  
XX  
XX WO200166752-A2.  
PN  
XX 13-SEP-2001.  
PD  
XX 07-MAR-2001; 2001WO-US07371.  
PF  
XX 07-MAR-2000; 2000US-0187518.  
PR  
XX 12-JAN-2001; 2001US-0261557.  
XX  
XX (WHEED ) WHITEHEAD INST BIOMEDICAL RES.  
PA  
XX Wang PJ, Page DC;  
PI  
XX WPI; 2001-570774/64.  
DR N-PSDB; AAS13646.  
XX  
XX Novel reproduction-specific protein, useful for treating disorders of  
PT reduced sperm count, enhancing/increasing sperm count and/or sperm  
PT activity -  
XX  
XX Claim 22; Fig 48; 151pp; English.  
PS

XX The present invention relates to the isolation of novel mammalian and  
CC human reproductive-specific proteins (AAU07859-AAU07899), and the  
CC nucleic acids encoding them. The nucleic acids encoding  
CC reproductive-specific proteins are useful for diagnosing infertility  
CC which is a result of reduced sperm count, reduced sperm motility,  
CC malformed sperm or combinations of these. The sequences of the invention  
CC are useful as markers for spermatogonial cells, for identifying genes or  
CC proteins characteristic of male infertility, diagnosing or aiding in  
CC the diagnosis of infertility in men, and for contraception in which  
CC sperm production or sperm count is reduced or defective sperm is  
CC produced. Antibodies to reproductive-specific proteins are useful for  
CC determining the presence of these proteins in a sample obtained from a  
CC man being assessed for infertility, for identifying the expression of  
CC genes in particular cell type or particular developmental stage, for  
CC studies of spermatogenesis, and for immunofluorescence of germ cells or  
CC in western blots for assessing the presence of the protein the antibody  
CC binds. The sequences of the invention are also useful for treating  
CC disorders of reduced sperm count, and for increasing sperm count and/or  
CC sperm activity. The nucleic acids of the invention are useful in gene  
CC therapy. AAU07859-AAU07882 represent the mammalian reproduction-specific  
CC proteins of the present invention.

XX Sequence 80 AA;

SQ  
Query Match 44.7%; Score 34; DB 22; Length 80;  
Best Local Similarity 50.0%; Pred. No. 51;  
Matches 7; Conservative 1; Mismatches 6; Indels 0; Gaps 0;

Qy 1 RYVDALNVLMMNI 14  
Db : | | | | |  
31 KYYSVLNVAFCNI 44

Search completed: February 17, 2004, 10:53:42  
Job time : 30.4158 secs

**THIS PAGE BLANK (USPTO)**



Result No.	Query		DB	ID	Description
	Score	Match			
1	39	38.6	78	16	Q98R59 mycoplasma
2	38	37.6	44	16	Q9NVK6 staphylococ
3	37	36.6	26	2	Q9F6Z5 staphylococ
4	37	36.6	37	16	Q9CLL0 yersinia pe
5	37	36.6	64	16	Q9EL85 oceanobacil
6	37	36.6	74	12	Q57258 vaccinia vi
7	37	36.6	75	16	Q9EHY5 shewanella
8	37	36.6	77	2	Q8KU65 enterococcu
9	37	36.6	87	12	Q9JFT1 ectromella
10	37	36.6	90	12	Q8QMP0 cowpox viru
11	37	36.6	95	2	O87254 lactococcus
12	37	36.6	97	16	Q9RF87 fusobacteri
13	37	36.6	100	5	Q8TOD7 drosophila
14	36.5	36.1	52	16	Q9EBS4 shewanella
15	36	35.6	29	2	O68162 staphylococ
16	36	35.6	36	2	Q8KYL2 bacillus an

```
AC Q8NVK6;
DT 01-OCT-2002 (TrEMBLrel. 22, Created)
DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
DT 01-OCT-2002 (TrEMBLrel. 22, Last annotation update)
DE Delta-hemolysin.
GN HLD OR MW1959.
OS Staphylococcus aureus (strain MW2).
OC Bacteria; Firmicutes; Bacillales; Staphylococcus.
OX NCBI_TaxID=196620;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=22040717; PubMed=12044378;
RA Baba T., Takeuchi F., Kuroda M., Yuzawa H., Aoki K.-I., Oguchi A.,
RA Nagai Y., Iwama N., Asano K., Naimi T., Kuroda H., Cui L.,
RA Yamamoto K., Hiramatsu K.;
RT "Genome and virulence determinants of high virulence community-
RT acquired MRSA"; 1827(2002).
RL Lancet 359:1819-1827(2002).
DR EMBL; AF004829; BAB95824.1; -.
KW Complete proteome.
SQ SEQUENCE 44 AA; 5039 MW; 24B1AG1E598DF6A8 CRC64;

Query Match 37.6%; Score 38; DB 16; Length 44;
Best Local Similarity 44.4%; Pred. No. 1.3e+02;
Matches 8; Conservative 3; Mismatches 7; Indels 0; Gaps 0;

QY 2 VLMAMNIIISKEKKEIKWI 19
Db 17 ISMAQDIISTISDLVKWI 34

RESULT 3
Q9F6Z5 PRELIMINARY; PRT; 26 AA.
AC Q9F6Z5;
DT 01-MAR-2001 (TrEMBLrel. 16, Created)
DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT 01-MAR-2001 (TrEMBLrel. 16, Last annotation update)
DE Delta hemolysin.
GN HLD.
OS Staphylococcus aureus.
OC Bacteria; Firmicutes; Bacillales; Staphylococcus.
OX NCBI_TaxID=1280;
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN=RN5881;
RA Jarraud S., Lyon G.J., Figueiredo A.M.S., Gerard L., Vandenesch F.,
RA Etienne J., Muir T.W., Novick R.P.;
RT "Exfoliatin-producing strains define a fourth agr specificity group in
RT Staphylococcus aureus.";
RL J. Bacteriol. 182:6517-6522(2000).
DR EMBL; AF288215; AAG03054.1; -.
DR HSPG; P01506; IDTC.
SQ SEQUENCE 26 AA; 3009 MW; 9F9BF3C5FE4DFD2 CRC64;

Query Match 36.6%; Score 37; DB 2; Length 26;
Best Local Similarity 50.0%; Pred. No. 1.2e+02;
Matches 8; Conservative 2; Mismatches 6; Indels 0; Gaps 0;

QY 4 MAMNIIISKEKKEIKWI 19
Db 1 MAQDIISTISDLVKWI 16

RESULT 4
Q8CLLO PRELIMINARY; PRT; 37 AA.
AC Q8CLLO;
DT 01-MAR-2003 (TrEMBLrel. 23, Created)
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE Hypothetical.
```

```
GN Y1112.
OS Yersinia pestis.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Yersinia.
OX NCBI_TaxID=632;
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN=KIMS / Biovar Mediaevalis;
RX MEDLINE=22137863; PubMed=12142430;
RA Deng W., Burland V., Plunkett G. III, Boutin A., Mayhew G.F., Liss P.,
RA Perna N.T., Rose D.J., Mau B., Zhou S., Schwartz D.C.,
RA Fetherston J.D., Lindler L.E., Brubaker R.R., Plano G.V.,
RA Straley S.C., McDonough K.A., Nilles M.L., Matson J.S., Blattner F.R.,
RA Perry R.D.;
RT "Genome sequence of Yersinia pestis KIM.";
RL J. Bacteriol. 184:4601-4611(2002).
DR EMBL; AB013714; AAM84690.1; -.
KW Hypothetical protein.
SQ SEQUENCE 37 AA; 4466 MW; ECDEF7F667957AB87 CRC64;

Query Match 36.6%; Score 37; DB 16; Length 37;
Best Local Similarity 40.0%; Pred. No. 1.6e+02;
Matches 6; Conservative 5; Mismatches 4; Indels 0; Gaps 0;

QY 4 MAMNIIISKEKKEIKW 18
Db 1 MSPNIIAEHDKILKW 15

RESULT 5
Q8EL85 PRELIMINARY; PRT; 64 AA.
AC Q8EL85;
DT 01-MAR-2003 (TrEMBLrel. 23, Created)
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE Hypothetical protein.
GN OB3346.
OS Oceanobacillus iheyensis.
OC Bacteria; Firmicutes; Bacillales; Oceanobacillus.
OX NCBI_TaxID=182710;
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN=HTE831 / DSM 14371 / JCM 11309;
RX MEDLINE=2220767; PubMed=1223376;
RA Takami H., Takaki Y., Uchiyama I.;
RT "Genome sequence of Oceanobacillus iheyensis isolated from the Iheya
RT Ridge and its unexpected adaptive capabilities to extreme
RT environments.";
RL Nucleic Acids Res. 30:3927-3935(2002).
DR EMBL; AP004604; BAC15302.1; -.
KW Hypothetical protein; Complete proteome.
SQ SEQUENCE 64 AA; 7353 MW; EB13F6E557963693 CRC64;

Query Match 36.6%; Score 37; DB 16; Length 64;
Best Local Similarity 50.0%; Pred. No. 2.6e+02;
Matches 7; Conservative 4; Mismatches 3; Indels 0; Gaps 0;

QY 2 VLMAMNIIISKEKKE 15
Db 49 VLMTYGIVAEKKE 62

RESULT 6
Q57258 PRELIMINARY; PRT; 74 AA.
AC Q57258;
DT 01-JUN-1998 (TrEMBLrel. 06, Created)
DT 01-JUN-1998 (TrEMBLrel. 06, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE Putative 8.5k protein.
GN MVA179R.
OS Vaccinia virus (strain Ankara).
```

```

OC Viruses; dsDNA viruses, no RNA stage; Poxviridae; Chordopoxvirinae;
OC Orthopoxvirus.
OX NCBI_TaxID=126794;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Ankara;
RA Antoine G., Scheiflinger F., Falkner F.G., Dörner F.;
RT "The complete genomic sequence of the Modified Vaccinia Ankara (MVA)
RT strain.";
RL Submitted (MAR-1997) to the EMBL/GenBank/DBJ databases.
DR EMBL: U94848; AAB96552.1; --
SQ SEQUENCE 74 AA; 8501 MW; 7152A06380829716 CRC64;

Query Match 36.6%; Score 37; DB 12; Length 74;
Best Local Similarity 56.2%; Pred. No. 3e+02;
Matches 9; Conservative 2; Mismatches 5; Indels 0; Gaps 0;

QY 1 NVLMANNIISKEKEIKI 16
Db 9 NVEDIMNEIDREKEI 24

RESULT 7
Q8EHYS PRELIMINARY; PRT; 75 AA.
AC Q8EHYS;
DT 01-MAR-2003 (TrEMBLrel. 23, Created)
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DE Conserved hypothetical protein.
GN S01069.
OS Shewanella oneidensis.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Alteromonadales;
OC Alteromonadaceae; Shewanella.
OX NCBI_TaxID=70863;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=MR-1;
RX MEDLINE=22297686; PubMed=12368813;
RA Heidelberg J.F., Paulsen I.T., Nelson K.E., Gaidos B.J., Nelson W.C.,
RA Read T.D., Eisen J.A., Seshadri R., Ward N., Methe B., Clayton R.A.,
RA Meyer T., Tsapin A., Scott J., Beaman M., Brinkac L., Daugherty S.,
RA DeBoy R.T., Dodson R.J., Durkin A.S., Haft D.H., Kolonay J.F.,
RA Madupu R., Peterson J.D., Unayam L.A., White O., Wolf A.M.,
RA Vamathevan J., Weidman J., Impraim M., Lee K., Berry K., Lee C.,
RA Mueller J., Khouri H., Gill J., Utterback T.R., McDonald L.A.,
RA Feldblyum T.V., Smith H.O., Venter J.C., Nealon K.H., Fraser C.M.;
RT "Genome sequence of the dissimilatory metal ion-reducing bacterium
RT Shewanella oneidensis.";
RL Nat. Biotechnol. 20:1118-1123(2002).
DR EMBL: AB015551; AAN54141.1; --
DR TIGR: S01069; --
KW Hypothetical protein; Complete proteome.
SQ SEQUENCE 75 AA; 8592 MW; D243608FA8177F2C CRC64;

Query Match 36.6%; Score 37; DB 16; Length 75;
Best Local Similarity 21.1%; Pred. No. 3e+02;
Matches 4; Conservative 12; Mismatches 3; Indels 0; Gaps 0;

QY 1 NVLMANNIISKEKEIKWI 19
Db 11 SVVVAISLWMSDIKKLRL 29

RESULT 8
Q8KU65 PRELIMINARY; PRT; 77 AA.
AC Q8KU65;
DT 01-OCT-2002 (TrEMBLrel. 22, Created)
DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE EF0096.
GN EF0096.

```

```

OS Enterococcus faecalis (Streptococcus faecalis).
OC Bacteria; Firmicutes; Lactobacillales; Enterococcaceae; Enterococcus.
OX NCBI_TaxID=1351;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=22061258; PubMed=12066186;
RA Shankar N., Baghdadyan A.S., Gilmore M.S.;
RT "Modulation of virulence within a pathogenicity island in vancomycin-
RT resistant Enterococcus faecalis.";
RL Nature 417:746-750(2002).
DR EMBL: AF454824; AAM75299.1; --
DR InterPro: IPR001367; HTH_DtXR.
DR Pfam: PF01325; Fe_dep_repress; 1.
SQ SEQUENCE 77 AA; 8884 MW; 00F76CDED10F5C47 CRC64;

Query Match 36.6%; Score 37; DB 2; Length 77;
Best Local Similarity 46.7%; Pred. No. 3.1e+02;
Matches 7; Conservative 5; Mismatches 3; Indels 0; Gaps 0;

QY 3 LMAMNIISKEKEIK 17
Db 13 LLAIQELSKKKLK 27

RESULT 9
Q9JFT1 PRELIMINARY; PRT; 87 AA.
AC Q9JFT1;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-OCT-2000 (TrEMBLrel. 15, Last annotation update)
DE C5R.
GN Ectromelia virus.
OS Viruses; dsDNA viruses, no RNA stage; Poxviridae; Chordopoxvirinae;
OC Orthopoxvirus.
OX NCBI_TaxID=12643;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Moscow;
RX MEDLINE=20192152; PubMed=10725549;
RA Chen N., Buller R.M.L., Wall E.M., Upton C.;
RT "Analysis of host response modifier ORFs of ectromelia virus, the
RT causative agent of mousepox.";
RL Virus Res. 66:155-173(2000).
DR EMBL: AF012825; AAC99564.1; --
SQ SEQUENCE 87 AA; 9879 MW; ECAC2FA1023BAACE CRC64;

Query Match 36.6%; Score 37; DB 12; Length 87;
Best Local Similarity 56.2%; Pred. No. 3.4e+02;
Matches 9; Conservative 2; Mismatches 5; Indels 0; Gaps 0;

QY 1 NVLMANNIISKEKEI 16
Db 22 NVEDIMNEIDREKEI 37

RESULT 10
Q8QMP0 PRELIMINARY; PRT; 90 AA.
AC Q8QMP0;
DT 01-JUN-2002 (TrEMBLrel. 21, Created)
DT 01-JUN-2002 (TrEMBLrel. 21, Last sequence update)
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE V196.
OS Cowpox virus (CPV).
OC Viruses; dsDNA viruses, no RNA stage; Poxviridae; Chordopoxvirinae;
OC Orthopoxvirus.
OX NCBI_TaxID=10243;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Brighton Red;
RX MEDLINE=83117629; PubMed=6961398;

```

SQ SEQUENCE 95 AA; 10967 MW; 2044EF16E363P911 CRC64;  
 Query Match 36.6%; Score 37; DB 2; Length 95;  
 Best Local Similarity 60.0%; Pred. No. 3.7e+02;  
 Matches 9; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 5 AMNIIISKEKKEIKWI 19  
      | | | | |  
 DB 14 AONTFEKKEDKII 28  
      | | | | |

RESULT 12  
 Q8RF87 PRELIMINARY; PRT; 97 AA.  
 ID Q8RF87  
 AC Q8RF87; 01-JUN-2002 (TREMBLrel. 21, Created)  
 DT 01-JUN-2002 (TREMBLrel. 21, Last sequence update)  
 DE 01-JUN-2002 (TREMBLrel. 21, Last annotation update)  
 DE Hypothetical protein FN0829.  
 GN FN0829.  
 OS Fusobacterium nucleatum (subsp. nucleatum).  
 OC Bacteria; Fusobacteria; Fusobacterales; Fusobacteriaceae;  
 OC Fusobacterium.  
 CX NCBI\_TaxID=76956;  
 RX [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=ATCC 25586;  
 RX MEDLINE=21986394; PubMed=11899109;  
 RA Kapralat V., Anderson I., Ivanova N., Reznik G., Los T., Lykidis A.,  
 RA Bhattacharyya A., Bartman A., Gardner W., Grechkin G., Zhu L.,  
 RA Vasileva O., Chu L., Kogan Y., Chaga O., Goldstein E., Bernal A.,  
 RA Larsen N., D'Souza M., Walunas T., Pusich G., Haseelkorn R.,  
 RA Fongstein M., Kyripides N., Overbeek R.;  
 RT "Genome sequence and analysis of the oral bacterium Fusobacterium  
 nucleatum strain ATCC 25586".  
 RL J. Bacteriol. 184:2005-2018(2002).  
 RL EMBL; AB010592; AAL95025.1; -;  
 KW Hypothetical protein; Complete proteome.  
 SQ SEQUENCE 97 AA; 11914 MW; DL6C48E41F9FD7EA CRC64;

Query Match 36.6%; Score 37; DB 16; Length 97;  
 Best Local Similarity 42.9%; Pred. No. 3.8e+02;  
 Matches 6; Conservative 7; Mismatches 1; Indels 0; Gaps 0;

QY 2 VLMMNNIISKKE 15  
      ::: : | | | |  
 DB 10 LLILSIFSQEKKE 23  
      ::: : | | | |

RESULT 13  
 Q8TD7 PRELIMINARY; PRT; 100 AA.  
 ID Q8TD7  
 AC Q8TD7; 01-JUN-2002 (TREMBlrel. 21, Created)  
 DT 01-JUN-2002 (TREMBlrel. 21, Last sequence update)  
 DE 01-MAR-2003 (TREMBlrel. 23, Last annotation update)  
 DE LD09564P.  
 GN SQD OR CG16901 OR CGI7791.  
 OS Drosophila melanogaster (Fruit fly).  
 OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;  
 OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;  
 OC Ephydroidea; Drosophilidae; Drosophila.  
 CX NCBI\_TaxID=7227;  
 RX [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=Berkely.  
 RA Stapleton M., Brokstein P., Hong L., Agbayani A., Carlson J.,  
 RA Champe M., Chavez C., Dorsett V., Faxian D., Frise E., George R.,  
 RA Gonzalez M., Guarin H., Li P., Liao G., Miranda A., Mungall C.J.,  
 RA Nunco J., Pacleb J., Paragas V., Park S., Phouanavong S., Wan K.,  
 RA Yu C., Lewis S.B., Rubin G.M., Celinker S.;  
 DR Submitted (DEC-2001) to the EMBL/Genbank/DDSI databases.  
 DR EMBL; AY069386; AAL39531.1; -.



```

RESULT 17
QBRYN3
ID QBRYN3 PRELIMINARY; PRT; 62 AA.
AC QBRYN3
DT 01-JUN-2002 (TrEMBLrel. 21, Created)
DT 01-JUN-2002 (TrEMBLrel. 21, Last sequence update)
DT 01-OCT-2002 (TrEMBLrel. 22, Last annotation update)
DE P0663E10.12 protein.
GN P0663E10.12
OS Oryza sativa. (japonica cultivar-group).
OC Eukaryota; Viridiplantae; Streptophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC Ehrhartoideae; Oryzaceae; Oryza.
OX NCBI_TaxID=39947;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=cv. Nipponbare;
RA Sasaki T., Matsumoto T., Yamamoto K.;
RT "Oryza sativa nipponbare (GA3) genomic DNA, chromosome 1, PAC
RT clone:P0663E10.1";
RL Submitted (NOV-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AP004317; BAB90729.1; -.
DR Gramene; QBRYN3; -.
SQ SEQUENCE 62 AA; 7432 MW; 8A2F84AD87B913EA CRC64;

Query Match 35.6%; Score 36; DB 10; Length 62;
Best Local Similarity 66.7%; Pred. No. 3.6e+02;
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 11 KKKKKKKKK 19
DB 32 KRKEEIEWI 40

RESULT 18
Q96Y48
ID Q96Y48 PRELIMINARY; PRT; 65 AA.
AC Q96Y48;
RX PubMed=11572479;
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE Hypothetical protein STS241.
GN STS241.
OS Sulfolobus tokodaii.
OC Archaea; Crenarchaeota; Thermoprotei; Sulfolobales; Sulfolobaceae;
OC Sulfolobus.
OX NCBI_TaxID=111955;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=JCM 10545 / 7;
RX PubMed=11572479;
RA Kawarabayashi Y., Hino Y., Horikawa H., Jin-no K., Takahashi M.,
RA Sekine M., Baba S.-I., Ankaï A., Kosugi H., Hosoyama A., Fukui S.,
RA Nagai Y., Nishijima K., Otsuka R., Nakazawa H., Takamiya M., Kato Y.,
RA Yoshizawa T., Tanaka T., Kudoh Y., Yamazaki J., Kushida N., Oguchi A.,
RA Aoki K.-I., Masuda S., Yanagii M., Nishimura M., Yamagishi A.,
RA Oshima T., Kikuchi H.;
RT "Complete genome sequence of an aerobic thermophilic
RT Crenarchaeon, Sulfolobus tokodaii strain7.";
RL DNA Res. 8:123-140(2001).
DR EMBL; AP000989; BAB67429.1; -.
KW Hypothetical protein; Complete proteome.
SQ SEQUENCE 65 AA; 7748 MW; 178C864B70EDCFB5 CRC64;

Query Match 35.6%; Score 36; DB 17; Length 65;
Best Local Similarity 53.8%; Pred. No. 3.8e+02;
Matches 7; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

QY 7 NIIKKKKKKKK 19
DB 13 NIVEKIKKREWI 25

RESULT 19
Q8EQ09
ID Q8EQ09 PRELIMINARY; PRT; 66 AA.
AC Q8EQ09;
DT 01-MAR-2003 (TrEMBLrel. 23, Created)
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE Hypothetical conserved protein.
GN O81920.
OS Oceanobacillus thelyensis.
OC Bacteria; Firmicutes; Bacillales; Oceanobacillus.
OX NCBI_TaxID=182710;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=HTE831 / DSM 14371 / JCM 11309;
RX MEDLINE=22220767; PubMed=12233376;
RA Takami H., Takaki Y., Uchiyama I.;
RT "Genome sequence of Oceanobacillus thelyensis isolated from the Iheya
RT Ridge and its unexpected adaptive capabilities to extreme
RT environments.";
RL Nucleic Acids Res. 30:3927-3935(2002).
DR EMBL; AP004599; BAC13876.1; -.
KW Hypothetical protein; Complete proteome.
SQ SEQUENCE 66 AA; 7775 MW; 4456C54A0BF9D73D CRC64;

Query Match 35.6%; Score 36; DB 16; Length 66;
Best Local Similarity 44.4%; Pred. No. 3.9e+02;
Matches 8; Conservative 3; Mismatches 7; Indels 0; Gaps 0;

QY 3 LMAMNIIKKKKKK 20
DB 49 LMAKLCVKKKKKK 66

RESULT 20
Q8V4T5
ID Q8V4T5 PRELIMINARY; PRT; 74 AA.
AC Q8V4T5;
DT 01-MAR-2002 (TrEMBLrel. 20, Created)
DT 01-MAR-2002 (TrEMBLrel. 20, Last sequence update)
DT 01-MAR-2002 (TrEMBLrel. 20, Last annotation update)
DE A44R.
GN A44R.
OS Monkeypox virus.
OC Viruses; dsDNA viruses, no RNA stage; Poxviridae; Chordopoxvirinae;
OC Orthopoxvirus.
OX NCBI_TaxID=10244;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Zaire-96-I-16;
RX MEDLINE=21592287; PubMed=11734207;
RA Shchelkunov S.N., Totmenin A.V., Babkin I.V., Safronov P.F.,
RA Ryazankina O.I., Petrov N.A., Gutorov V.V., Uvarova E.A.,
RA Mikheev M.V., Sisler J.R., Esposito J.J., Jahrling P.B., Moss B.,
RA Sandakchchiev L.S.;
RT "Human monkeypox and smallpox viruses: genomic comparison.";
RL FEBS Lett. 509:66-70(2001).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=Zaire-96-I-16;
RX Shchelkunov S.N., Totmenin A.V., Safronov P.F., Gutorov V.V.,
RA Ryazankina O.I., Petrov N.A., Babkin I.V., Uvarova E.A.,
RA Esposito J.J., Moss B., Sisler J.R., Jahrling P.B., Sandakchchiev L.S.;
RT Submitted (MAY-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF380138; AAL40612.1; -.
SQ SEQUENCE 74 AA; 8376 MW; CC9232CB8408A8A9 CRC64;

Query Match 35.6%; Score 36; DB 12; Length 74;
Best Local Similarity 35.3%; Pred. No. 4.3e+02;
Matches 6; Conservative 7; Mismatches 4; Indels 0; Gaps 0;

```

QY 1 NVLMANNIISKEKEIK 17  
 DB 38 DVLLAQSVAVEARDVK 54

## RESULT 21

ID Q8V2L4 PRELIMINARY; PRT; 78 AA.  
 AC Q8V2L4;  
 DT 01-MAR-2002 (TrEMBLrel. 20, Created)  
 DT 01-MAR-2002 (TrEMBLrel. 20, Last sequence update)  
 DT 01-OCT-2002 (TrEMBLrel. 22, Last annotation update)  
 DE Hypothetical 8.8 kDa protein (CMP160R).  
 GN CMP160R.  
 OS Camelpox virus (strain CP-1).  
 OC Viruses; dsDNA viruses, no RNA stage; Poxviridae; Chordopoxvirinae;  
 OC Orthopoxvirus.  
 OX NCBI\_TaxID=203174;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=M-96;  
 RA Afonso C.L., Tulman E.R., Lu Z., Zeak L., Zaitsev V.L.,  
 RA Kerembekova U.Z., Sandybaev N.T., Kutish G.F., Rock D.L.;  
 RT "The genome of camelpox virus.";  
 RL Submitted (Oct-2001) to the EMBL/GenBank/DBJ databases.  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=CMS;  
 RX PubMed=11907336;  
 RA Gubser C., Smith G.L.;  
 RT "The sequence of camelpox virus shows it is most closely related to  
 RT variola virus, the cause of smallpox.";  
 RL J. Gen. Virol. 83:855-872(2002).  
 RN [3]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=CMS;  
 RA Gubser C., Smith G.L.;  
 RL Submitted (Oct-2000) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; AF438165; AAL73870.1; -  
 DR EMBL; AY009089; AAG37659.1; -  
 KW Hypothetical protein.  
 SQ SEQUENCE 78 AA; 8808 MW; CCB88471E95A3E6D CRC64;

Query Match 35.6%; Score 36; DB 12; Length 78;  
 Best Local Similarity 35.3%; Pred. No. 4.5e+02;  
 Matches 6; Conservative 7; Mismatches 4; Indels 0; Gaps 0;

QY 1 NVLMANNIISKEKEIK 17  
 DB 42 DVLLAQSVAVEARDVK 58

## RESULT 22

ID Q89663 PRELIMINARY; PRT; 78 AA.  
 AC Q89663;  
 DT 01-NOV-1996 (TrEMBLrel. 01, Created)  
 DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)  
 DT 01-OCT-2002 (TrEMBLrel. 22, Last annotation update)  
 DE 8.7K protein.  
 GN SALFGR OR MVAL56R.  
 OS Vaccinia virus, and  
 OS Vaccinia virus (strain Ankara).  
 OC Viruses; dsDNA viruses, no RNA stage; Poxviridae; Chordopoxvirinae;  
 OC Orthopoxvirus.  
 OX NCBI\_TaxID=10245, 126794;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=WR;  
 RX MEDLINE=69203625; PubMed=4891218;  
 RA Ueda Y., Ito M., Tagaya I.;  
 RT "A specific surface antigen induced by poxvirus.";  
 RL Virology 38:180-182(1969).

RN [2]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=WR;  
 RX MEDLINE=80052057; PubMed=501796;  
 RA Payne L.G.;  
 RT "Identification of the vaccinia hemagglutinin polypeptide from a cell  
 RT system yielding large amounts of extracellular enveloped virus.";  
 RL J. Virol. 31:147-155(1979).  
 RN [3]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=WR;  
 RX MEDLINE=86181588; PubMed=3008418;  
 RA Shida H.;  
 RT "Nucleotide sequence of the vaccinia virus hemagglutinin gene.";  
 RL Virology 150:451-462(1986).  
 RN [4]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=WR;  
 RX MEDLINE=90094307; PubMed=26000076;  
 RA Traktman P., Anderson M.K., Rempel R.E.;  
 RT "Vaccinia virus encodes an essential gene with strong homology to  
 RT protein kinases.";  
 RL J. Biol. Chem. 264:21458-21461(1989).  
 RN [5]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=WR;  
 RX MEDLINE=89381686; PubMed=2778436;  
 RA Smith G.L., Howard S.T., Chan Y.S.;  
 RT "Vaccinia virus encodes a family of genes with homology to serine  
 RT proteinase inhibitors.";  
 RL J. Gen. Virol. 70:2333-2343(1989).  
 RN [6]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=WR;  
 RX MEDLINE=90111697; PubMed=2607336;  
 RA Howard S.T., Smith G.L.;  
 RT "Two early vaccinia virus genes encode polypeptides related to protein  
 RT kinases.";  
 RL J. Gen. Virol. 70:3187-3201(1989).  
 RN [7]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=WR;  
 RX MEDLINE=89094985; PubMed=2783466;  
 RA Kotwal G.J., Moss B.;  
 RT "Vaccinia virus encodes two proteins that are structurally related to  
 RT members of the plasma serine protease inhibitor superfamily.";  
 RL J. Virol. 63:600-606(1989).  
 RN [8]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=WR;  
 RX MEDLINE=90016845; PubMed=2552411;  
 RA Smith G.L., de Carlos A., Chan Y.S.;  
 RT "Vaccinia virus encodes a thymidylate kinase gene: sequence and  
 RT transcriptional mapping.";  
 RL Nucleic Acids Res. 17:7581-7590(1989).  
 RN [9]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=WR;  
 RX MEDLINE=90067908; PubMed=2587253;  
 RA Kerr S.M., Smith G.L.;  
 RT "Vaccinia virus encodes a polypeptide with DNA ligase activity.";  
 RL Nucleic Acids Res. 17:9039-9050(1989).  
 RN [10]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=WR;  
 RX MEDLINE=90067909; PubMed=2555782;  
 RA Smith G.L., Chan Y.S., Kerr S.M.;  
 RT "Transcriptional mapping and nucleotide sequence of a vaccinia virus  
 RT gene encoding a polypeptide with extensive homology to DNA ligases.";  
 RL Nucleic Acids Res. 17:9051-9062(1989).  
 RN [11]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=WR;

Tue Feb 17 11:55:51 2004

RX MEDLINE=91011369; PubMed=2170578;  
RA Zhou J., Crawford L., McLean L., Sun X.Y., Stanley M., Almond N.,  
RA Smith G.L.;  
RT "Increased antibody responses to human papillomavirus type 16 L1  
RT protein expressed by recombinant vaccinia virus lacking serine  
RT protease inhibitor genes.";  
RL J. Gen. Virol. 71:2185-2190(1990).  
RN [12]  
RP SEQUENCE FROM N.A.  
RC STRAIN=WR;  
RX MEDLINE=90320131; PubMed=2196742;  
RA Ueda Y., Morikawa S., Matsura Y.;  
RT "Identification and nucleotide sequence of the gene encoding a surface  
RT antigen induced by vaccinia virus.";  
RL Virology 177:588-594(1990).  
RN [13]  
RP SEQUENCE FROM N.A.  
RC STRAIN=WR;  
RX MEDLINE=91021027; PubMed=2219722;  
RA Goebel S.J., Johnson G.P., Perkus M.E., Davis S.W., Winslow J.P.,  
RA Paolletti E.;  
RT "The complete DNA sequence of vaccinia virus.";  
RL Virology 179:247-266(1990).  
RN [14]  
RP SEQUENCE FROM N.A.  
RC STRAIN=WR;  
RX MEDLINE=91021029; PubMed=2219723;  
RA Colinas R.J., Goebel S.J., Davis S.W., Johnson G.P., Norton E.K.,  
RA Paolletti E.;  
RT "A DNA ligase gene in the Copenhagen strain of vaccinia virus is  
RT nonessential for viral replication and recombination.";  
RL Virology 179:267-275(1990).  
RN [15]  
RP SEQUENCE FROM N.A.  
RC STRAIN=WR;  
RX MEDLINE=92097554; PubMed=1756739;  
RA Kerr S.M., Johnston L.H., Odell M., Duncan S.A., Law K.M., Smith G.L.;  
RT "Vaccinia DNA ligase complements Saccharomyces cerevisiae cdc9  
RT localizes in cytoplasmic factories and affects virulence and virus  
RT sensitivity to DNA damaging agents.";  
RL EMBO J. 10:4343-4350(1991).  
RN [16]  
RP SEQUENCE FROM N.A.  
RC STRAIN=WR;  
RX MEDLINE=92041833; PubMed=1657913;  
RA Hughes S.J., Johnston L.H., de Carlos A., Smith G.L.;  
RT "Vaccinia virus encodes an active thymidylate kinase that complements  
RT a cdc8 mutant of Saccharomyces cerevisiae.";  
RL J. Biol. Chem. 266:20103-20109(1991).  
RN [17]  
RP SEQUENCE FROM N.A.  
RC STRAIN=WR;  
RX MEDLINE=91170931; PubMed=1826022;  
RA Smith G.L., Chan Y.S.;  
RT "Two vaccinia virus proteins structurally related to the interleukin-1  
RT receptor and the immunoglobulin superfamily.";  
RL J. Gen. Virol. 72:511-518(1991).  
RN [18]  
RP SEQUENCE FROM N.A.  
RC STRAIN=WR;  
RX MEDLINE=91259063; PubMed=2045793;  
RA Smith G.L., Chan Y.S., Howard S.T.;  
RT "Nucleotide sequence of 42 kbp of vaccinia virus strain WR from near  
RT the right inverted terminal repeat.";  
RL J. Gen. Virol. 72:1349-1376(1991).  
RN [19]  
RP SEQUENCE FROM N.A.  
RC STRAIN=WR;  
RX MEDLINE=91332999; PubMed=1870190;  
RA Blasco R., Cole N.B., Moss B.;  
RT "Sequence analysis, expression, and deletion of a vaccinia virus gene  
RT encoding a homolog of profilin, a eukaryotic actin-binding protein.";  
RL J. Virol. 65:4598-4608(1991).

RN [20]  
RP SEQUENCE FROM N.A.  
RC STRAIN=WR;  
RX MEDLINE=91111981; PubMed=1989387;  
RA Kerr S.M., Smith G.L.;  
RT "Vaccinia virus DNA ligase is nonessential for virus replication;  
RT recovery of plasmids from virus-infected cells.";  
RL Virology 180:625-632(1991).  
RN [21]  
RP SEQUENCE FROM N.A.  
RC STRAIN=WR;  
RX MEDLINE=91111982; PubMed=1846491;  
RA Howard S.T., Chan Y.S., Smith G.L.;  
RT "Vaccinia virus homologues of the Shope fibroma virus inverted  
RT terminal repeat proteins and a discontinuous ORF related to the tumor  
RT necrosis factor receptor family.";  
RL Virology 180:633-647(1991).  
RN [22]  
RP SEQUENCE FROM N.A.  
RC STRAIN=WR;  
RX MEDLINE=91134977; PubMed=1994573;  
RA Takahashi-Nishimaki F., Funahashi S., Miki K., Hashizume S.,  
RA Sugimoto M.;  
RT "Regulation of plaque size and host range by a vaccinia virus gene  
RT related to complement system proteins.";  
RL Virology 181:158-164(1991).  
RN [23]  
RP SEQUENCE FROM N.A.  
RC STRAIN=WR;  
RX MEDLINE=93008237; PubMed=1339315;  
RA Spriggs M.K., Hruby D.E., Maliszewski C.R., Pickup D.J., Sims J.E.,  
RA Buller R.M., Vanslyke J.;  
RT "Vaccinia and cowpox viruses encode a novel secreted interleukin-1  
RT binding protein.";  
RL Cell 71:145-152(1992).  
QY 1 NVLMANNIIISKEKIK 17  
Db 42 DVLLAQSVAVEAKDKV 58  
Query Match 35.6%; Score 36; DB 12; Length 78;  
Best Local Similarity 35.3%; Pred. No. 4.5e+02;  
Matches 6; Conservative 7; Mismatches 4; Indels 0; Gaps 0;  
ID Q8Y9P9 PRELIMINARY; PRT; 80 AA.  
AC Q8Y9P9;  
DT 01-MAR-2002 (TREMBLrel. 20, Created)  
DT 01-MAR-2002 (TREMBLrel. 20, Last sequence update)  
DT 01-MAR-2002 (TREMBLrel. 20, Last annotation update)  
DE Hypothetical protein lmo0475.  
GN LMO0475.  
OS Listeria monocytogenes.  
OC Bacteria; Firmicutes; Bacillales; Listeriaceae; Listeria.  
OX NCBI\_TaxID=1639;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=EGD-e / Serovar 1/2a;  
RX MEDLINE=21537279; PubMed=11679669;  
RA Glaser P., Frangeul L., Buchrieser C., Ruaniok C., Amend A.,  
RA Baquero F., Berche P., Bloecher H., Brandt P., Chakraborty T.,  
RA Charbit A., Chetouani F., Couve E., de Daruvar A., Dehoux P.,  
RA Domann E., Dominguez-Bernal G., Duchaud E., Durant L., Duesurget O.,  
RA Entian K.-D., Faihi H., Garcia-del Portillo F., Garrido P.,  
RA Gautier L., Goebel W., Gomez-Lopez N., Hain T., Hauf J., Jackson D.,  
RA Jones L.-M., Kaerst U., Kref J., Kuhn M., Kunst F., Kurapkat G.,  
RA Madueno E., Maitournam A., Mata Vicente J., Ng E., Nedjari H.,  
RA Nordsiek G., Novella S., de Pablo B., Perez-Diaz J.-C., Purcell R.,  
RA Remmel B., Rose M., Schluter T., Simoes N., Tierrez A.,  
RA Vasquez-Boland J.-A., Voss H., Wehland J., Cossart P.;  
RT "Comparative genomics of Listeria species."



```
RL Science 294:849-852(2001).
DR EMBL; AL591975; CAC98554.1; -.
DR Listlist; IM000475; -.
KW Hypothetical protein; Complete proteome.
SQ SEQUENCE 80 AA; 9128 MW; E3FC5C721686CFC6 CRC64;

Query Match 35.6%; Score 36; DB 16; Length 80;
Best Local Similarity 40.0%; Pred. No. 4.6e+02;
Matches 6; Conservative 5; Mismatches 4; Indels 0; Gaps 0;

QY 4 MAMNIISKEKEIKW 18
   :|:|:|:|:|:|
Db 46 LATSIVNVKKEYLW 60

RESULT 24
Q8FEY1 PRELIMINARY; PRT; 88 AA.
ID Q8FEY1
AC Q8FEY1;
DT 01-MAR-2003 (TrEMBLrel. 23, Created)
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DE Hypothetical protein ydfk.
GN YDFK OR C3145.
OS Escherichia coli O6.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Escherichia.
OX NCBI_TaxID=217992;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=06:HI / CFT073 / ATCC 700928;
RX MEDLINE=22388234; PubMed12471157;
RA Welch R.A., Burland V., Plunkett G. III, Redford P., Roesch P.,
RA Rasko D., Buckles E.L., Liou S.-R., Boutin A., Hackett J., Stroud D.,
RA Mayhew G.F., Rose D.J., Zhou S., Schwartz D.C., Perna N.T.,
RA Mobley H.L.T., Donnenberg M.S., Blattner F.R.;
RT "Extensive mosaic structure revealed by the complete genome sequence
RT of uropathogenic Escherichia coli."
RL Proc. Natl. Acad. Sci. U.S.A. 99:17020-17024(2002).
DR EMBL; AE016764; AAN81595.1; -.
KW Hypothetical protein; Complete proteome.
SQ SEQUENCE 88 AA; 10092 MW; 76D61A2F2869094A CRC64;

Query Match 35.6%; Score 36; DB 16; Length 88;
Best Local Similarity 44.4%; Pred. No. 5e+02;
Matches 8; Conservative 2; Mismatches 6; Indels 2; Gaps 1;

QY 1 NVLMAMNIISKEKEIKW 18
   :|:|:|:|:|:|
Db 5 NIL--RNIFMKSRTLKW 20

RESULT 25
Q8VTS1 PRELIMINARY; PRT; 90 AA.
ID Q8VTS1
AC Q8VTS1;
DT 01-MAR-2002 (TrEMBLrel. 20, Created)
DT 01-MAR-2002 (TrEMBLrel. 20, Last sequence update)
DE Hypothetical protein ydfk.
GN YDFK OR C3145.
OS Listeria welshimeri.
OC Bacteria; Firmicutes; Bacillales; Listeriaceae; Listeria.
OX NCBI_TaxID=1643;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CF3VP;
RA Cai S., Wiedmann M.;
RT "Characterization of the prfa virulence gene cluster insertion site in
RT non-hemolytic Listeria spp.: Probing the evolution of the Listeria
RT virulence gene island."
RL Submitted (NOV-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF322005; AAL36603.1; -.

RL Science 294:849-852(2001).
DR EMBL; AL591975; CAC98554.1; -.
DR Listlist; IM000475; -.
KW Hypothetical protein; Complete proteome.
SQ SEQUENCE 80 AA; 9128 MW; E3FC5C721686CFC6 CRC64;

Query Match 35.6%; Score 36; DB 16; Length 80;
Best Local Similarity 40.0%; Pred. No. 4.6e+02;
Matches 6; Conservative 5; Mismatches 4; Indels 0; Gaps 0;

QY 4 MAMNIISKEKEIKW 18
   :|:|:|:|:|:|
Db 46 LATSIVNVKKEYLW 60

RESULT 24
Q8FEY1 PRELIMINARY; PRT; 88 AA.
ID Q8FEY1
AC Q8FEY1;
DT 01-MAR-2003 (TrEMBLrel. 23, Created)
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DE Hypothetical protein ydfk.
GN YDFK OR C3145.
OS Escherichia coli O6.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Escherichia.
OX NCBI_TaxID=217992;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=06:HI / CFT073 / ATCC 700928;
RX MEDLINE=22388234; PubMed12471157;
RA Welch R.A., Burland V., Plunkett G. III, Redford P., Roesch P.,
RA Rasko D., Buckles E.L., Liou S.-R., Boutin A., Hackett J., Stroud D.,
RA Mayhew G.F., Rose D.J., Zhou S., Schwartz D.C., Perna N.T.,
RA Mobley H.L.T., Donnenberg M.S., Blattner F.R.;
RT "Extensive mosaic structure revealed by the complete genome sequence
RT of uropathogenic Escherichia coli."
RL Proc. Natl. Acad. Sci. U.S.A. 99:17020-17024(2002).
DR EMBL; AE016764; AAN81595.1; -.
KW Hypothetical protein; Complete proteome.
SQ SEQUENCE 88 AA; 10092 MW; 76D61A2F2869094A CRC64;

Query Match 35.6%; Score 36; DB 16; Length 88;
Best Local Similarity 44.4%; Pred. No. 5e+02;
Matches 8; Conservative 2; Mismatches 6; Indels 2; Gaps 1;

QY 1 NVLMAMNIISKEKEIKW 18
   :|:|:|:|:|:|
Db 5 NIL--RNIFMKSRTLKW 20

RESULT 25
Q8VTS1 PRELIMINARY; PRT; 90 AA.
ID Q8VTS1
AC Q8VTS1;
DT 01-MAR-2002 (TrEMBLrel. 20, Created)
DT 01-MAR-2002 (TrEMBLrel. 20, Last sequence update)
DE Hypothetical protein ydfk.
GN YDFK OR C3145.
OS Listeria welshimeri.
OC Bacteria; Firmicutes; Bacillales; Listeriaceae; Listeria.
OX NCBI_TaxID=1643;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CF3VP;
RA Cai S., Wiedmann M.;
RT "Characterization of the prfa virulence gene cluster insertion site in
RT non-hemolytic Listeria spp.: Probing the evolution of the Listeria
RT virulence gene island."
RL Submitted (NOV-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF322005; AAL36603.1; -.

DR InterPro; IPR001236; ldh.
DR Pfam; PF02866; ldh_C; 1.
KW Oxidoreductase.
FT NON_TER 1
SQ SEQUENCE 90 AA; 9917 MW; 033D966F56B0B413 CRC64;

Query Match 35.6%; Score 36; DB 2; Length 90;
Best Local Similarity 40.0%; Pred. No. 5.1e+02;
Matches 6; Conservative 6; Mismatches 3; Indels 0; Gaps 0;

QY 3 LMAMNIISKEKEIK 17
   :|:|:|:|:|:|
Db 58 IVENLNDRKEQMK 72

RESULT 26
Q8VTS6 PRELIMINARY; PRT; 91 AA.
ID Q8VTS6
AC Q8VTS6;
DT 01-MAR-2002 (TrEMBLrel. 20, Created)
DT 01-MAR-2002 (TrEMBLrel. 20, Last sequence update)
DE Lactate dehydrogenase (Fragment).
GN LDH.
OS Listeria innocua.
OC Bacteria; Firmicutes; Bacillales; Listeriaceae; Listeria.
OX NCBI_TaxID=1642;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=dd644;
RA Cai S., Wiedmann M.;
RT "Characterization of the prfa virulence gene cluster insertion site in
RT non-hemolytic Listeria spp.: Probing the evolution of the Listeria
RT virulence gene island."
RL Submitted (NOV-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF322004; AAL36598.1; -.
DR InterPro; IPR001236; ldh.
DR Pfam; PF02866; ldh_C; 1.
KW Oxidoreductase.
FT NON_TER 1
SQ SEQUENCE 91 AA; 10030 MW; D07DDEE31E74385B CRC64;

Query Match 35.6%; Score 36; DB 2; Length 91;
Best Local Similarity 40.0%; Pred. No. 5.1e+02;
Matches 6; Conservative 6; Mismatches 3; Indels 0; Gaps 0;

QY 3 LMAMNIISKEKEIK 17
   :|:|:|:|:|:|
Db 59 IVENLNDRKEQMK 73

RESULT 27
Q9RIG5 PRELIMINARY; PRT; 65 AA.
ID Q9RIG5
AC Q9RIG5;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-JUN-2000 (TrEMBLrel. 21, Last annotation update)
DE Hypothetical 7.4 kDa protein (Fragment).
OS Yersinia enterocolitica.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Yersinia.
OX NCBI_TaxID=630;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=99426800; PubMed=10496882;
RA Bach S., Buchrieser C., Prentice M., Guiry A., Meadek T.,
RA Carniel E.;
RT "The high-pathogenicity island of Yersinia enterocolitica YE8081
RT undergoes low-frequency deletion but not precise excision, suggesting
RT recent stabilization in the genome."
RL Infect. Immun. 67:5091-5099(1999).
DR EMBL; AJ238284; CAB57394.1; -.

```

SQ SEQUENCE 51 AA; 5581 MW; 1C8B9EED80CDB59 CRC64;

Query Match 34.7%; Score 35; DB 4; Length 51;  
Best Local Similarity 70.0%; Pred. No. 4.4e+02;  
Matches 7; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 6 MNIISKEKE 15  
DB 37 MNQLGKEKE 46

RESULT 30  
Q8NOT8 PRELIMINARY; PRT; 54 AA.

ID Q8NOT8  
AC Q8NOT8  
DT 01-OCT-2002 (TRENBLrel. 22, Created)  
DT 01-OCT-2002 (TRENBLrel. 22, Last sequence update)  
DT 01-MAR-2003 (TRENBLrel. 23, Last annotation update)  
DE Similar to hypothetical gene LOC133874.  
OS Homo sapiens (Human)  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
OC NCBI\_TaxID=9606;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC TISSUE=Brain;  
RA Submitted (MAY-2002) to the EMBL/GenBank/DBJ databases.  
RL EMBL; BC030767; AAH30767.1;  
DR InterPro; IPR002453; Beta\_tubulin.  
DR PROSITE; PS00228; TUBULIN\_B\_AUTOREG; 1.  
KW Hypothetical protein.  
SQ SEQUENCE 54 AA; 6227 MW; F4EF7945303F09BE CRC64;

Query Match 34.7%; Score 35; DB 4; Length 54;  
Best Local Similarity 50.0%; Pred. No. 4.6e+02;  
Matches 8; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

QY 2 VLMAMNIISKEKEIK 17  
DB 37 VIKNINTISSELKIK 52

RESULT 31  
Q8XLC3 PRELIMINARY; PRT; 57 AA.

ID Q8XLC3  
AC Q8XLC3  
DT 01-MAR-2002 (TRENBLrel. 20, Created)  
DT 01-MAR-2002 (TRENBLrel. 20, Last sequence update)  
DT 01-MAR-2002 (TRENBLrel. 20, Last annotation update)  
DE Hypothetical protein CPE119.  
GN CPE119.  
OS Clostridium perfringens.  
OC Bacteria; Firmicutes; Clostridia; Clostridiales; Clostridiaceae;  
OC Clostridium.  
OX NCBI\_TaxID=1502;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=13 / Type A;  
RC PubMed=11792842;  
RA Shimizu T., Ohtani K., Hirakawa H., Ohshima K., Yamashita A.,  
RA Shiba T., Ogasawara N., Hattori M., Kuhara S., Hayashi H.,  
RT "Complete genome sequence of Clostridium perfringens, an anaerobic  
flesh-eater."  
RL Proc. Natl. Acad. Sci. U.S.A. 99:996-1001(2002).  
DR EMBL; AP003189; BAB80825.1;  
KW Hypothetical protein; Complete proteome.  
SQ SEQUENCE 57 AA; 6854 MW; B46DF46C56CE59DD CRC64;

Query Match 34.7%; Score 35; DB 16; Length 57;  
Best Local Similarity 58.3%; Pred. No. 4.8e+02;  
Matches 7; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

SQ SEQUENCE 65 AA; 7404 MW; 90E14C97C1216DF6 CRC64;

Query Match 35.1%; Score 35.5; DB 2; Length 65;  
Best Local Similarity 47.1%; Pred. No. 4.5e+02;  
Matches 8; Conservative 6; Mismatches 2; Indels 1; Gaps 1;

QY 1 NVL-MAMNIISKEKEI 16  
DB 14 NVMFIAIGVIAKEKEHI 30

RESULT 28  
O57370 PRELIMINARY; PRT; 40 AA.

ID O57370  
AC O57370  
DT 01-JUN-1998 (TRENBLrel. 06, Created)  
DT 01-JUN-1998 (TRENBLrel. 06, Last sequence update)  
DT 01-MAR-2003 (TRENBLrel. 23, Last annotation update)  
DE Hoxc9 protein (Fragment).  
GN HOCX9A OR HOCX9  
OS Brachydanio rerio (Zebrafish) (Danio rerio).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes;  
OC Cyprinidae; Danio.  
OC NCBI\_TaxID=7955;  
RN [1]  
RP SEQUENCE FROM N.A.  
RA Prince V.E., Joly L., Ekker M., Ho R.K.;  
RT "Zebrafish hox genes: genomic organization and modified colinear  
expression patterns in the trunk."  
RL Submitted (AUG-1997) to the EMBL/GenBank/DBJ databases.  
DR EMBL; Y14540; CAA74875.1;  
DR ZFIN; ZDB-GENE-000328-5; hoxc9a.  
DR InterPro; IPR001356; Homeobox.  
DR Pfam; PF00046; homeobox; 1.  
DR SMART; SM00389; HOX; 1.  
DR PROSITE; PS00027; HOMEBOX\_1; 1.  
DR PROSITE; PS00071; HOMEBOX\_2; 1.  
FT NON TER  
SQ SEQUENCE 40 AA; 5097 MW; AD06E7BD625901FF CRC64;

Query Match 34.7%; Score 35; DB 13; Length 40;  
Best Local Similarity 47.1%; Pred. No. 3.5e+02;  
Matches 8; Conservative 3; Mismatches 6; Indels 0; Gaps 0;

QY 1 NVL-MAMNIISKEKEIK 17  
DB 22 NRRMKNKNKNDKSK 38

RESULT 29  
Q96PS2 PRELIMINARY; PRT; 51 AA.

ID Q96PS2  
AC Q96PS2  
DT 01-DEC-2001 (TRENBLrel. 19, Created)  
DT 01-DEC-2001 (TRENBLrel. 19, Last sequence update)  
DT 01-DEC-2001 (TRENBLrel. 19, Last annotation update)  
DE FGF-2 activity-associated protein 3.  
GN GAF3.  
OS Homo sapiens (Human)  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
OC NCBI\_TaxID=9606;  
RN [1]  
RP SEQUENCE FROM N.A.  
RA Gan Y., Au J.L.S., Lu J., Wientjes G.M.;  
RT "The gene associated with FGF-2 activity-3 (GAF3)."  
RL Submitted (DEC-1999) to the EMBL/GenBank/DBJ databases.  
DR EMBL; AF220235; AAL11916.1;

```
QY 7 NIISKKEIKW 18
DB 3 NIILKLNKILW 14

RESULT 32
Q8C1Q0 PRELIMINARY; PRT; 81 AA.
AC Q8C1Q0;
DT 01-MAR-2003 (TrEMBLrel. 23, Created)
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE Calmodulin-related protein.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Small intestine;
RX MEDLINE=22354683; PubMed=12466851;
RA The FANTOM Consortium.
RA The RIKEN Genome Exploration Research Group Phase I & II Team;
RT "Analysis of the mouse transcriptome based on functional annotation of
RT 60,770 full-length cDNAs."
RL Nature 420:563-573 (2002).
DR EMBL; AK008039; BAC25195.1; -
SQ SEQUENCE 81 AA; 9496 MW; 0FB651BC92082A6 CRC64;

Query Match 34.7%; Score 35; DB 11; Length 81;
Best Local Similarity 53.8%; Pred. No. 6.6e+02;
Matches 7; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

QY 2 VLMAMNIISKEIK 14
DB 14 ILLMLMADKEIK 26

RESULT 33
Q95JP8 PRELIMINARY; PRT; 83 AA.
AC Q95JP8;
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE Hypothetical 9.5 kDa protein.
OS Macaca fascicularis (Crab eating macaque) (Cynomolgus monkey).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Cercopithecoidea;
OC Cercopithecinae; Macaca.
OX NCBI_TaxID=9541;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Testis;
RA Hashimoto K., Osada N., Hida M., Kusuda J., Tanuma R., Hirai M.,
RA Terao K., Sugano S.;
RT "Isolation of novel full-length cDNA clones from macaque testis cDNA
RT libraries."
RL Submitted (AUG-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AB070133; BAB63078.1; -
KW Hypothetical protein.
SQ SEQUENCE 83 AA; 9450 MW; 7EC2FFA2706CA3F1 CRC64;

Query Match 34.7%; Score 35; DB 6; Length 83;
Best Local Similarity 35.3%; Pred. No. 6.8e+02;
Matches 6; Conservative 6; Mismatches 5; Indels 0; Gaps 0;

QY 4 MAMNIISKEIKWIG 20
DB 9 IAQGLKSSQKEIKLVG 25

RESULT 34
Q44637 PRELIMINARY; PRT; 85 AA.
AC Q44637;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE Arginyl-tRNA synthetase (Fragment).
GN ARG5.
OS Buchnera aphidicola.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Buchnera.
OX NCBI_TaxID=9;
RN [1]
RP SEQUENCE FROM N.A.
RA Munson M.A., Baumann L., Baumann P.;
RT "Buchnera aphidicola (a prokaryotic endosymbiont of aphids) contains a
RT putative 16S rRNA operon unlinked to the 23S rRNA-encoding gene;
RT sequence determination and promoter and terminator analysis."
RL Submitted (AUG-1993) to the EMBL/GenBank/DBJ databases.
DR EMBL; L18928; AAA72379.1; -
DR InterPro; IPR005148; N.
DR Pfam; PF03485; N-Arg; 1.
KW Aminoacyl-tRNA synthetase.
FT NON_TER 85
SQ SEQUENCE 85 AA; 9982 MW; 162FCA5A447FAA68 CRC64;

Query Match 34.7%; Score 35; DB 2; Length 85;
Best Local Similarity 58.3%; Pred. No. 6.9e+02;
Matches 7; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 6 MNIIISKEIK 17
DB 1 MNILSKLNKIK 12

RESULT 35
Q97XM8 PRELIMINARY; PRT; 88 AA.
AC Q97XM8;
DT 01-OCT-2001 (TrEMBLrel. 18, Created)
DT 01-OCT-2001 (TrEMBLrel. 18, Last sequence update)
DT 01-OCT-2001 (TrEMBLrel. 18, Last annotation update)
DE Partial transposase ISC1190.
GN SSO8687.
OS Sulfolobus solfataricus.
OC Archaea; Crenarchaeota; Thermoprotei; Sulfolobales; Sulfolobaceae;
OC Sulfolobus.
OX NCBI_TaxID=2287;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC 35092 / DSM 1617 / P2;
RX MEDLINE=21332296; PubMed=11427726;
RA She Q., Singh R.K., Confalonieri F., Zivanovic Y., Allard G.,
RA Aways M.J., Chan-Welher C.C.-Y., Clausen I.G., Curtis B.A.,
RA De Moors A., Brauso G., Fletcher C., Gordon P.M.K.,
RA Heikamp-de Jong I., Jeffries A.C., Kozera C.J., Medina N., Peng X.,
RA Thi-Ngoc H.P., Redder P., Schenk M.E., Theriault C., Tolstrup N.,
RA Charlebois R.L., Doolittle W.F., Duguet M., Gaasterland T.,
RA Garrett R.A., Ragan M.A., Sensen C.W., Van der Oost J.;
RT "The complete genome of the crenarchaeon Sulfolobus solfataricus P2."
RL Proc. Natl. Acad. Sci. U.S.A. 98:7835-7840 (2001).
DR EMBL; AE006782; AAK41895.1; -
KW Complete proteome.
SQ SEQUENCE 88 AA; 10466 MW; PFFEC9DA6554CDEA CRC64;

Query Match 34.7%; Score 35; DB 17; Length 88;
Best Local Similarity 53.3%; Pred. No. 7.2e+02;
Matches 8; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 3 LMAMNIISKEIK 17
DB 48 LMKQGFIEKNDYK 62
```

```
RESULT 36
Q8H247 PRELIMINARY; PRT; 94 AA.
AC Q8H247;
DT 01-MAR-2003 (TrEMBLrel. 23, Created)
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE Anthocyanidin synthase (Fragment).
GN ANS.
OS Vaccinium myrtillus.
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
OC Asteridae; Ericaceae; Vaccinioideae; Vacciniaceae; Vaccinium.
OX NCBI_TaxID=180763;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=22264001; PubMed=12376640;
RA Jaakola L., Maatta K., Pirttila A.M., Torronen R., Karenlampi S.,
RA Hohtola A.;
RT "Expression of Genes Involved in Anthocyanin Biosynthesis in Relation
RT to Anthocyanin, Proanthocyanidin, and Flavonol Levels during Bilberry
RT Fruit Development.";
RL Plant Physiol. 130:729-739 (2002).
DR EMBL, AY123768; AAM96895.1; -.
FT NON_TER 1
FT TER 94
SQ SEQUENCE 94 AA; 10548 MW; DDD96D7D62CF9923 CRC64;

Query Match 34.7%; Score 35; DB 10; Length 94;
Best Local Similarity 31.2%; Pred. No. 7.6e+02;
Matches 5; Conservative 5; Mismatches 6; Indels 0; Gaps 0;

QY 3 LMANNIIS-KEKKEIKW 18
DB 76 ILHRLVNEKVRISW 91

RESULT 37
Q8YRS4 PRELIMINARY; PRT; 60 AA.
AC Q8YRS4;
DT 01-MAR-2002 (TrEMBLrel. 20, Created)
DT 01-MAR-2002 (TrEMBLrel. 20, Last sequence update)
DT 01-MAR-2002 (TrEMBLrel. 20, Last annotation update)
DE Hypothetical protein Asr3369.
GN ASR3369.
OS Anabaena sp. (strain PCC 7120).
OC Bacteria; Cyanobacteria; Nostocales; Nostocaceae; Nostoc.
OX NCBI_TaxID=103690;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=21595285; PubMed=11759840;
RA Kaneko T., Nakamura Y., Wolk C.P., Kuritz T., Sasamoto S.,
RA Watanabe A., Iriiguchi M., Ishikawa A., Kawashima K., Kimura T.,
RA Kishida Y., Kohara M., Matsumoto M., Matsuno A., Muraki A.,
RA Nakazaki N., Shimo S., Sugimoto M., Takazawa M., Yamada M.,
RA Yasuda M., Tabata S.;
RT "Complete genomic sequence of the filamentous nitrogen-fixing
RT cyanobacterium Anabaena sp. strain PCC 7120.";
RL DNA Res. 8:205-213 (2001).
DR EMBL, AP003592; BAB75068.1; -.
KW Hypothetical protein; Complete proteome.
SQ SEQUENCE 60 AA; 6701 MW; B26CFA037642BC94 CRC64;

Query Match 34.2%; Score 34.5; DB 16; Length 60;
Best Local Similarity 44.4%; Pred. No. 6.1e+02;
Matches 8; Conservative 5; Mismatches 4; Indels 1; Gaps 1;

QY 2 VLMAMNIIS-KEKKEIKW 18
DB 17 ILIAPNIEIKERRNQKW 34

RESULT 38
Q9RPN3 PRELIMINARY; PRT; 90 AA.
AC Q9RPN3;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE ORF4.
GN ORF4.
OS Yersinia pestis.
OG Plasmid pYC.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Yersinia.
OX NCBI_TaxID=632;
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN=CH971662;
RA Dong X.Q., Lindler L.E., Chu M.C.;
RT "Complete DNA sequence and analysis of an emerging cryptic plasmid
RT isolated from Yersinia pestis.";
RL Plasmid 43:144-148 (2000).
DR EMBL, AF152923; AAF05100.1; -.
KW Plasmid.
SQ SEQUENCE 90 AA; 10070 MW; 80D1C30322A92D7C CRC64;

Query Match 34.2%; Score 34.5; DB 2; Length 90;
Best Local Similarity 40.0%; Pred. No. 8.8e+02;
Matches 8; Conservative 5; Mismatches 4; Indels 3; Gaps 1;

QY 2 VLMAMNIIS-KEKKEIKW 18
DB 10 VLMALISLGGCEETKSADW 29

RESULT 39
Q8E4H3 PRELIMINARY; PRT; 90 AA.
AC Q8E4H3;
DT 01-MAR-2003 (TrEMBLrel. 23, Created)
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE 30S ribosomal protein S16.
GN RPS16 OR GBS1428.
OS Streptococcus agalactiae (serotype III).
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC Streptococcus.
OX NCBI_TaxID=216495;
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN=NEM316 / Serotype III;
RA Glaser P., Rusniok C., Buchrieser C., Chevalier F., Frangeul L.,
RA Maadok T., Zouine M., Couve E., Lalioui L., Poyart C., Trieu-Cuot P.,
RA Kunst F.;
RT "Genome sequence of Streptococcus agalactiae, a pathogen causing
RT invasive neonatal disease.";
RL Mol. Microbiol. 45:1499-1513 (2002).
DR EMBL, AL766850; CAD47087.1; -.
DR Sagaliet, gbs1428; -.
KW Ribosomal protein; Complete proteome.
SQ SEQUENCE 90 AA; 10282 MW; CDCA763D78007CBE CRC64;

Query Match 34.2%; Score 34.5; DB 16; Length 90;
Best Local Similarity 35.0%; Pred. No. 8.8e+02;
Matches 7; Conservative 8; Mismatches 4; Indels 1; Gaps 1;

QY 1 NVLMAMNIIS-KEKKEIKW 19
DB 41 NPLVAENQVTKERSVLEWL 60
```

```
RESULT 40
Q8DIW5 Q8DIW5 PRELIMINARY; PRT; 90 AA.
AC Q8DIW5;
DT 01-MAR-2003 (TrEMBLrel. 23, Created)
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE Ribosomal protein S16.
GN RSP OR SAG1358.
OS Streptococcus agalactiae (serotype V).
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC Streptococcus.
OX NCBI_TaxID=216466;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=2603 V/R / Serotype V;
RX MEDLINE=22222988; PubMed=12200547;
RA Tettelin H., Maignani V., Cieslewicz M.J., Eisen J.A., Peterson S.,
RA Wessels L.C., Wolf A.M., Beanan M.J., Brinkac L.M., Daugherty S.C.,
RA Madoff R.T., Durkin A.S., Kolonay J.F., Madupu R., Lewis M.R.,
RA Radune D., Fedorova N.B., Scanlan D., Khouri H., Mulligan S.,
RA Carty H.A., Cline R.T., Van Aken S.E., Gill J., Scarselli M., Mora M.,
RA Tacchini E.T., Brettoni C., Galli G., Mariani M., Vegni F., Maione D.,
RA Rinaudo D., Rappuoli R., Telford J.D., Kasper D.L., Grandi G.,
RA Fraser C.M.;
RT "Complete genome sequence and comparative genomic analysis of an
RT emerging human pathogen, serotype V Streptococcus agalactiae.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:12391-12396(2002).
DR EMBL; AE014254; AA00229.1; -.
DR TIGR; SAG1358; -.
KW Complete proteome.
SQ SEQUENCE 90 AA; 10282 MW; CDCA763D78007CBE CRC64;

Query Match 34.2%; Score 34.5; DB 16; Length 90;
Best Local Similarity 35.0%; Pred. No. 8.e+02;
Matches 7; Conservative 8; Mismatches 4; Indels 1; Gaps 1;

QY 1 NVLMAMNIIIS-KEKKEIKWI 19
Db 41 NPLVAENQVTKIEKRVLEWL 60

RESULT 41
Q8VS57 Q8VS57 PRELIMINARY; PRT; 98 AA.
AC Q8VS57;
DT 01-MAR-2002 (TrEMBLrel. 20, Created)
DT 01-MAR-2002 (TrEMBLrel. 20, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE GyrA (Fragment).
GN GYRA.
OS Proteus mirabilis.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Proteus.
OX NCBI_TaxID=584;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC 29906;
RA Weigel L.M., Anderson G.J., Tenover F.C.;
RT "gyrA and parC Mutations Associated with Fluoroquinolone Resistance in
RT Proteus mirabilis.";
RL Submitted (JUL-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF397169; AAL57340.1; -.
DR InterPro; IPR002205; DNA_topoisoIV.
DR Pfam; PF00521; DNA_topoisoIV; 1.
DR ProDom; PD000742; DNA_topoisoIV; 1.
DR SMART; SM00434; TOP4c; 1.
FT NON_TER 1
FT NON_TER 98
SQ SEQUENCE 98 AA; 10963 MW; 052CD25D0FASF1F CRC64;

Query Match 34.2%; Score 34.5; DB 2; Length 98;

Best Local Similarity 33.3%; Pred. No. 9.5e+02;
Matches 8; Conservative 5; Mismatches 6; Indels 5; Gaps 1;

QY 2 VLMAMNIIISKE-----KKEIKWIG 20
Db 8 VLFAMNVILGNDWKNPKYKKSARVVG 31

RESULT 42
Q8CNM6 Q8CNM6 PRELIMINARY; PRT; 25 AA.
AC Q8CNM6;
DT 01-MAR-2003 (TrEMBLrel. 23, Created)
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE Delta-haemolysin.
GN SE1634.
OS Staphylococcus epidermidis.
OC Bacteria; Firmicutes; Bacillales; Staphylococcus.
OX NCBI_TaxID=1282;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC 12228;
RA Zhang Y., Ren S., Li H., Fu G., Lu L., Lu G., Jia J., Tu Y., Qin Z.,
RA Chen Z., Wen Y.;
RL Submitted (NOV-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AE016749; AA005233.1; -.
KW Complete proteome.
SQ SEQUENCE 25 AA; 2820 MW; 342155FE4DFCF19 CRC64;

Query Match 33.7%; Score 34; DB 16; Length 25;
Best Local Similarity 50.0%; Pred. No. 3.3e+02;
Matches 8; Conservative 2; Mismatches 6; Indels 0; Gaps 0;

QY 4 MAMNIIISKEKKEIKWI 19
Db 1 MAADIISTIGDLVKWI 16

RESULT 43
Q9SD80 Q9SD80 PRELIMINARY; PRT; 54 AA.
AC Q9SD80;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-MAY-2000 (TrEMBLrel. 13, Last annotation update)
DE Hypothetical 6.1 kDa protein.
GN F13G24.240.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eurosids II; Brassicales; Brassicaceae; Arabidopsis.
OX NCBI_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.
RA Bevan M., Van Der Schueren J., Chuang Y.J., Voet M., Robben J.,
RA Volckaert G., Bancroft I., Mewes H.W., Lemcke K., Mayer K.F.X.;
RL Submitted (DEC-1999) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RA EU Arabidopsis sequencing project;
RL Submitted (DEC-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL; AL133421; CAB62616.1; -.
KW Hypothetical protein.
SQ SEQUENCE 54 AA; 6091 MW; 8DB3B77A2F0FC3A1 CRC64;

Query Match 33.7%; Score 34; DB 10; Length 54;
Best Local Similarity 40.0%; Pred. No. 6.6e+02;
Matches 6; Conservative 3; Mismatches 6; Indels 0; Gaps 0;

QY 4 MAMNIIISKEKKEIKW 18
Db 1 MVNVVVSIEKWKALW 15
```

Tue Feb 17 11:55:51 2004

RL Science 294:849-852(2001).  
DR EMBL: AL591983; CAD00510.1; -.  
DW Listlist; LMO2432; -  
KW Hypothetical protein; Complete proteome.  
SQ SEQUENCE 54 AA; 6310 MW; 94C858AE0304279E CRC64;  
  
Query Match 33.7%; Score 34; DB 16; Length 54;  
Best Local Similarity 50.0%; Pred. No. 6.6e+02;  
Matches 6; Conservative 3; Mismatches 3; Indels 0; Gaps 0;  
  
QY 6 MNIISKEKEIK 17  
:|:|:|:|:|  
Db 42 LNIIVDKKPKNIK 53  
  
Search completed: February 17, 2004, 10:56:21  
Job time : 30.3168 secs

RESULT 44  
Q928K6 PRELIMINARY; PRT; 54 AA.  
AC Q928K6; 01-DEC-2001 (TrEMBLrel. 19, Created)  
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)  
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)  
DE Hypothetical protein lin2526.  
GN LIN2526  
OS Listeria innocua.  
OC Bacteria; Firmicutes; Bacillales; Listeriaceae; Listeria.  
OX NCBI\_TaxID=1642;  
RN [1]  
RC STRAIN=CLIP 11262 / Serovar 6a;  
RX PubMed=11679669;  
RA Glaser P., Frangeul L., Buchrieser C., Rusniok C., Amend A.,  
BAguero F., Berche P., Bloecker H., Brandt P., Chakraborty T.,  
RA Charbit A., Chetouani F., Couve E., de Daruvar A., Dehoux P.,  
RA Domann E., Dominguez-Bernal G., Duchaud E., Durant L., Dussurget O.,  
RA Entian K.-D., Fsihi H., Garcia-del Portillo F., Garrido P.,  
RA Gautier L., Goebel W., Gomez-Lopez N., Hain T., Hauf J., Jackson D.,  
RA Jones L.-M., Kaerst U., Krest J., Kuhn M., Kunst F., Kurapkat G.,  
RA Madueno E., Maitournam A., Mata Vicente J., Ng E., Nedjari H.,  
RA Nordsiek G., Novella S., de Pablo B., Perez-Diaz J.-C., Purcell R.,  
RA Remmel B., Rose M., Schlueter T., Simoes N., Tierrez A.,  
RA Vazquez-Boland J.-A., Voss H., Wehland J., Cossart P.;  
RT "Comparative genomics of Listeria species.";  
RL Science 294:849-852(2001).  
DR EMBL: AL596172; CAC97753.1; -.  
DR Listlist; LMO2526; -.  
KW Hypothetical protein; Complete proteome.  
SQ SEQUENCE 54 AA; 6263 MW; 20C858BFE4B4508E CRC64;  
  
Query Match 33.7%; Score 34; DB 16; Length 54;  
Best Local Similarity 50.0%; Pred. No. 6.6e+02;  
Matches 6; Conservative 3; Mismatches 3; Indels 0; Gaps 0;  
  
QY 6 MNIISKEKEIK 17  
:|:|:|:|:|  
Db 42 LNIIVDKKPKNIK 53  
  
RESULT 45  
Q8Y4K6 PRELIMINARY; PRT; 54 AA.  
AC Q8Y4K6;  
DT 01-MAR-2002 (TrEMBLrel. 20, Created)  
DT 01-MAR-2002 (TrEMBLrel. 20, Last sequence update)  
DT 01-MAR-2002 (TrEMBLrel. 20, Last annotation update)  
DE Hypothetical protein lmo2432.  
GN LMO2432  
OS Listeria monocytogenes.  
OC Bacteria; Firmicutes; Bacillales; Listeriaceae; Listeria.  
OX NCBI\_TaxID=1639;  
RN [1]  
RC STRAIN=EGD-e / Serovar 1/2a;  
RX MEDLINE=21537279; PubMed=11679669;  
RA Glaser P., Frangeul L., Buchrieser C., Rusniok C., Amend A.,  
BAguero F., Berche P., Bloecker H., Brandt P., Chakraborty T.,  
RA Charbit A., Chetouani F., Couve E., de Daruvar A., Dehoux P.,  
RA Domann E., Dominguez-Bernal G., Duchaud E., Durant L., Dussurget O.,  
RA Entian K.-D., Fsihi H., Garcia-del Portillo F., Garrido P.,  
RA Gautier L., Goebel W., Gomez-Lopez N., Hain T., Hauf J., Jackson D.,  
RA Jones L.-M., Kaerst U., Krest J., Kuhn M., Kunst F., Kurapkat G.,  
RA Madueno E., Maitournam A., Mata Vicente J., Ng E., Nedjari H.,  
RA Nordsiek G., Novella S., de Pablo B., Perez-Diaz J.-C., Purcell R.,  
RA Remmel B., Rose M., Schlueter T., Simoes N., Tierrez A.,  
RA Vazquez-Boland J.-A., Voss H., Wehland J., Cossart P.;  
RT "Comparative genomics of Listeria species.";

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: February 17, 2004, 10:50:13 ; Search time 12.0792 Seconds  
(without alignments)  
159.230 Million cell updates/sec

Title: US-09-900-147-4

Perfect score: 101

Sequence: 1 NVLMANNIISKKEIKWIG 20

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283308 seqs, 96168682 residues

Total number of hits satisfying chosen parameters: 37673

Minimum DB seq length: 0

Maximum DB seq length: 100

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : PIR 76:\*

1: pir1:\*

2: pir2:\*

3: pir3:\*

4: pir4:\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	39	38.6	78	2 G90530	hypothetical prote
2	37	36.6	74	2 T37447	probable 8.5K prot
3	37	36.6	95	2 T43125	hypothetical prote
4	36	35.6	42	2 T07291	hypothetical prote
5	36	35.6	44	2 A89995	delta-hemolysin [1
6	36	35.6	64	2 D69406	hypothetical prote
7	36	35.6	78	2 JQ1781	SalpFR 8.8K protei
8	36	35.6	80	2 AD1134	hypothetical prote
9	36	35.6	88	2 B64888	ynae protein - Esc
10	36	35.6	88	2 C64909	yafk protein - Esc
11	35	34.7	26	1 LESAD	delta hemolysin -
12	35	34.7	74	2 G69286	hypothetical prote
13	35	34.7	88	2 H90329	partial transposas
14	34.5	34.2	60	2 AB2227	hypothetical prote
15	34	33.7	25	2 A33058	B52b delta-like to
16	34	33.7	54	2 T45629	hypothetical prote
17	34	33.7	54	2 AH1378	hypothetical prote
18	34	33.7	54	2 A11747	hypothetical prote
19	34	33.7	67	2 H97845	hypothetical prote
20	34	33.7	72	2 AE3378	hypothetical prote
21	34	33.7	67	2 S34404	ribosomal protein
22	34	33.7	79	2 AC1217	hypothetical prote
23	34	33.7	90	2 B81102	conserved hypotet
24	34	33.7	91	1 C64069	virulence-associat
25	34	33.7	95	2 F72864	AcOrf-117 protein
26	34	33.7	95	2 T41855	AcMNPV orf117 - Bo
27	34	33.7	96	2 B84142	hypothetical prote
28	34	33.7	97	2 C45681	hypothetical lik p
29	33.5	33.2	73	2 H71893	probable molybdopt

30	33.5	33.2	74	2 A64620	molybdopterin conv
31	33	32.7	33	2 E84146	hypothetical prote
32	33	32.7	42	2 S38633	ribosomal protein
33	33	32.7	46	2 B90196	probable SSV1 viru
34	33	32.7	61	2 I59337	mammary transformat
35	33	32.7	66	2 T43096	hypothetical prote
36	33	32.7	67	2 T12860	hypothetical prote
37	33	32.7	67	2 F97798	hypothetical prote
38	33	32.7	68	2 H90086	hypothetical prote
39	33	32.7	70	2 B71673	cold shock-like pr
40	33	32.7	70	2 E97827	cold shock-like pr
41	33	32.7	72	2 JQ1805	BlrR protein - vac
42	33	32.7	76	2 H90332	hypothetical prote
43	33	32.7	82	2 C90115	small nuclear ribo
44	33	32.7	87	2 F81321	hypothetical prote
45	33	32.7	88	2 A42527	BlrR protein - vac

ALIGNMENTS

RESULT 1

G90530

hypothetical protein MYPV\_1510 [imported] - Mycoplasma pulmonis (strain UAB CTIP)  
C;Species: Mycoplasma pulmonis  
C;Date: 24-May-2001 #sequence\_revision 24-May-2001 #text\_change 03-Aug-2001  
C;Accession: G90530

R;Chambaud, I.; Heilig, R.; Ferris, S.; Barbe, V.; Samson, D.; Gallisson, F.; Moszer, I.;

Nucleic Acids Res. 29, 2145-2153, 2001

A;Title: The complete genome sequence of the murine respiratory pathogen Mycoplasma pulmo

A;Reference number: A99512; MUID:21267165; PMID:11353084

A;Accession: G90530

A;Status: preliminary

A;Molecule type: DNA

A;Residues: 1-78 <KUR>

A;Cross-references: GB:AL445566; PID:gl04089564; PIDN:CAC13324.1; GSPDB:GN00153

A;Experimental source: strain UAB CTIP

C;Genetics:

A;Gene: MYPV\_1510

A;Genetic code: SGC3

Query Match 38.6%; Score 39; DB 2; Length 78;  
Best Local Similarity 42.1%; Pred. No. 49;  
Matches 8; Conservative 5; Mismatches 6; Indels 0; Gaps 0;

Qy 1 NVLMANNIISKKEIKW 19

Db 10 NVKLILDHISKPKREWI 28

RESULT 2

T37447

probable 8.5K protein - vaccinia virus (strain Ankara)  
C;Species: vaccinia virus  
C;Variety: strain Ankara

C;Date: 21-Jan-2000 #sequence\_revision 21-Jan-2000 #text\_change 21-Jan-2000

C;Accession: T37447

R;Antoine, G.; Scheiflinger, F.; Falkner, F.G.; Dorner, F.

submitted to the EMBL Data Library, March 1997

A;Description: The complete genomic sequence of the Modified Vaccinia Ankara (MVA) strain

A;Reference number: Z20877

A;Accession: T37447

A;Status: preliminary; translated from GB/EMBL/DBJ

A;Molecule type: DNA

A;Residues: 1-74 <ANT>

A;Cross-references: EMBL:U94848; PIDN:AAB96552.1

A;Experimental source: strain Ankara

C;Genetics:

A;Note: MVA179R

Query Match 36.6%; Score 37; DB 2; Length 74;  
Best Local Similarity 56.2%; Pred. No. 95;  
Matches 9; Conservative 2; Mismatches 5; Indels 0; Gaps 0;

QY 1 NVLMAMNIIISKEKEIKWI 16  
||| :|||  
Db 9 NVEDIMNEIDREKEEI 24

RESULT 3  
T43125  
hypothetical protein - Lactococcus lactis plasmid pMRC01  
C:Species: Lactococcus lactis  
C>Date: 02-Jun-2000 #sequence\_revision 02-Jun-2000 #text\_change 02-Jun-2000  
C:Accession: T43125  
R:Dougherty, B.A.; Hill, C.; Weidman, J.F.; Richardson, D.R.; Venter, J.C.; Ross, R.P.  
Mol. Microbiol. 29, 1029-1038, 1998  
A:Title: Sequence and analysis of the 60 kb conjugative, bacteriocin-producing plasmid pMRC01  
A:Reference number: Z22314  
A:Accession: T43125  
A:Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: DNA  
A:Residues: 1-95 <DOU>  
A:Cross-references: EMBL:AB001272; PIDN:AA056046.1  
A:Experimental source: strain DPC3147  
A:Genome: plasmid pMRC01  
A>Note: ORF00054

Query Match 36.6%; Score 37; DB 2; Length 95;  
Best Local Similarity 60.0%; Pred. No. 1.2e+02;  
Matches 9; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 5 AMNIIISKEKEIKWI 19  
| | | | | | | | | |  
Db 14 AONTFEKEKEDKII 28

RESULT 4  
T07291  
hypothetical protein 42c - Chlorella vulgaris chloroplast  
C:Species: Chloroplast Chlorella vulgaris  
C>Date: 14-May-1999 #sequence\_revision 14-May-1999 #text\_change 21-Jul-2000  
C:Accession: T07291  
R:Wakasugi, T.; Nagai, T.; Kapoor, M.; Sugita, M.; Ito, M.; Ito, S.; Taudzuki, J.; Nakae  
Proc. Natl. Acad. Sci. U.S.A. 94, 5967-5972, 1997  
A:Title: Complete nucleotide sequence of the chloroplast genome from the green alga Chlorella vulgaris  
A:Reference number: Z15985; MUID:97303241; PMID:9159184  
A:Accession: T07291  
A:Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: DNA  
A:Residues: 1-42 <WAK>  
A:Cross-references: EMBL:AB001684; NID:G2224352; PIDN:BA057939.1; PID:G2224455  
C:Genetics:  
A:Genome: chloroplast  
C:Keywords: chloroplast

Query Match 35.6%; Score 36; DB 2; Length 42;  
Best Local Similarity 47.1%; Pred. No. 75;  
Matches 8; Conservative 3; Mismatches 6; Indels 0; Gaps 0;

QY 3 LMAMNIIISKEKEIKWI 19  
| | | | | | | | | |  
Db 16 LFLIPFFSKKKKEIKQV 32

RESULT 5  
A89995  
delta-hemolysin [imported] - Staphylococcus aureus (strain N315)  
C:Species: Staphylococcus aureus  
C>Date: 10-May-2001 #sequence\_revision 10-May-2001 #text\_change 22-Oct-2001  
C:Accession: A89995  
R:Kuroda, M.; Ohta, T.; Uchiyama, I.; Baba, T.; Yuzawa, H.; Kobayashi, I.; Cui, L.; Oguc  
ma, A.; Mizutani-U, Y.; Kobayashi, N.; Sawano, T.; Inoue, R.; Kaito, C.; Sekimizu, K.;  
C.; Shiba, T.; Hattori, M.; Ogasawara, N.; Hayashi, H.; Hiratsatsu, K.  
Lancet 357, 1225-1240, 2001

A:Title: Whole genome sequencing of meticillin-resistant Staphylococcus aureus.  
A:Reference number: A89758; MUID:21311952; PMID:11418146  
A:Accession: A89995  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-44 <KUR>  
A:Cross-references: GB:BA000018; PID:913701829; PIDN:BA043122.1; GSPDB:GN00149  
A:Experimental source: strain N315  
C:Genetics:  
A:Gene: hld

Query Match 35.6%; Score 36; DB 2; Length 44;  
Best Local Similarity 44.4%; Pred. No. 79;  
Matches 8; Conservative 3; Mismatches 7; Indels 0; Gaps 0;

QY 2 VLMAMNIIISKEKEIKWI 19  
: | | | | | | | | | |  
Db 17 ISMAQDIISTIGDLVKWI 34

RESULT 6  
D69406  
hypothetical protein AF1253 - Archaeoglobus fulgidus  
C:Species: Archaeoglobus fulgidus  
C>Date: 05-Dec-1997 #sequence\_revision 05-Dec-1997 #text\_change 22-Oct-1999  
C:Accession: D69406  
R:Klenk, H.P.; Clayton, R.A.; Tomb, J.F.; White, O.; Nelson, K.E.; Ketchum, K.A.; Dodson,  
J.; Fleischmann, R.D.; Quackenbush, J.; Lee, N.H.; Sutton, G.G.; Gill, S.; Kirkness, E.F.;  
Glodek, A.; Zhou, L.; Overbeek, R.; Gocayne, J.D.; Weidman, J.F.; McDonald, L.  
Nature 390, 364-370, 1997  
A:Authors: Overbeek, T.; Cotton, M.D.; Spriggs, T.; Artiach, P.; Kaine, B.P.; Sykes, S.H.  
Smith, H.O.; Woese, C.R.; Venter, J.C.  
A:Title: The complete genome sequence of the hyperthermophilic, sulfate-reducing archaeon  
A:Reference number: A69250; MUID:98049343; PMID:9389475  
A:Accession: D69406  
A:Status: preliminary; nucleic acid sequence not shown; translation not shown  
A:Molecule type: DNA  
A:Residues: 1-64 <KLE>  
A:Cross-references: GB:AE001018; GB:AE000782; NID:G2689341; PIDN:AB090006.1; PID:G2649344

Query Match 35.6%; Score 36; DB 2; Length 64;  
Best Local Similarity 60.0%; Pred. No. 1.2e+02;  
Matches 9; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

QY 5 AMNIIISKEKEIKWI 19  
| | | | | | | | | |  
Db 47 AVEEITKANKEIKKI 61

RESULT 7  
JQ1781  
Salp6R 8.8K protein - vaccinia virus (strain WR and Ankara)  
C:Species: vaccinia virus  
C>Date: 30-Sep-1993 #sequence\_revision 30-Sep-1993 #text\_change 18-Feb-2000  
C:Accession: JQ1781; D40897; T37429 S.T.  
R:Smith, G.L.; Chan, Y.S.; Howard, S.T.  
J. Gen. Virol. 72, 1349-1376, 1991  
A:Title: Nucleotide sequence of 42kbp of vaccinia virus strain WR from near the right in  
A:Reference number: JQ1767; MUID:91259063; PMID:2045793  
A:Accession: JQ1781  
A:Molecule type: DNA  
A:Residues: 1-78 <SMI>  
A:Cross-references: DDBJ:D11079; NID:G222717; PIDN:BA01817.1; PID:d1002293; PID:G222732  
R:Blasco, R.; Cole, N.B.; Moss, B.  
J. Virol. 65, 4598-4608, 1991  
A:Title: Sequence analysis, expression, and deletion of a vaccinia virus gene encoding a  
A:Reference number: A40897; MUID:91332999; PMID:1870190  
A:Accession: D40897  
A:Molecule type: DNA  
A:Residues: 1-78 <BLA>  
A:Cross-references: GB:M72474; NID:G335761; PIDN:AAA48310.1; PID:G335765  
R:Antoine, G.; Scheiflinger, F.; Falkner, F.G.; Dörner, F.  
submitted to the EMBL Data Library, March 1997





F:/Modified site: N-formylmethionine #status experimental

Query Match 34.7%; Score 35; DB 1; Length 26;  
Best Local Similarity 50.0%; Pred. No. 66;  
Matches 8; Conservative 2; Mismatches 6; Indels 0; Gaps 0;

QY 4 MAMNIIISKEKEIKWI 19  
||:||||:||||  
Db 1 MAQDIISTIGDLVKWI 16

RESULT 12

G69286  
hypothetical protein AF0295 - Archaeoglobus fulgidus  
C/Species: Archaeoglobus fulgidus  
C/Date: 05-Dec-1997 #sequence\_revision 05-Dec-1997 #text\_change 22-Oct-1999

R/Klenk, H.P.; Clayton, R.A.; Tomb, J.F.; White, O.; Nelson, K.E.; Ketchum, K.A.; Dodson  
.; Fleischmann, R.D.; Quackenbush, J.; Lee, N.H.; Sutton, G.G.; Gill, S.; Kirkness, E.F.  
Glodek, A.; Zhou, L.; Overbeek, R.; Gocayne, J.D.; Weidman, J.F.; McDonald, L.  
Nature 390, 364-370, 1997  
A/Authors: Utterback, T.; Cotton, M.D.; Spriggs, T.; Artiach, P.; Kaine, B.P.; Sykes, S.  
Smith, H.O.; Woese, C.R.; Venter, J.C.  
A/Title: The complete genome sequence of the hyperthermophilic, sulfate-reducing archaeo  
A/Reference number: A69250; MUID:98049343; PMID:9389475  
A/Accession: G69286  
A/Status: preliminary; nucleic acid sequence not shown; translation not shown  
A/Molecule type: DNA  
A/Residues: 1-74 <KLE>  
A/Cross-references: GB:AE001084; GB:AE000782; NID:g2689407; PIDN:AAB90939.1; PID:g265034

Query Match 34.7%; Score 35; DB 2; Length 74;  
Best Local Similarity 52.9%; Pred. No. 28+02;  
Matches 9; Conservative 1; Mismatches 7; Indels 0; Gaps 0;

QY 2 VLMAMNIIISKEKEIKW 18  
||:||||:||||  
Db 5 VLEAGKIISPNEKVIW 21

RESULT 13

H90329  
partial transposase ISC1190 [imported] - Sulfolobus solfataricus  
C/Species: Sulfolobus solfataricus  
C/Date: 24-May-2001 #sequence\_revision 24-May-2001 #text\_change 24-May-2001

R/She, Q.; Singh, R.K.; Confalonieri, F.; Zivanovic, Y.; Allard, G.; Awayez, M.J.; Chan-  
Jong, I.; Jeffries, A.C.; Kozera, C.J.; Medina, N.; Peng, X.; Thi-Ngoc, H.P.; Redder, F.  
arrett, R.A.; Ragan, M.A.; Sengen, C.W.; Van der Oost, J.  
submitted to GenBank, April 2001  
A/Description: Sulfolobus solfataricus complete genome.  
A/Reference number: A99139  
A/Accession: H90329  
A/Status: preliminary  
A/Molecule type: DNA  
A/Residues: 1-88 <KUR>  
A/Cross-references: GB:AE006641; NID:gl3814935; PIDN:AAK41895.1; GSPDB:GN00155  
C/Genetics:  
A/Gene: SS08687

Query Match 34.7%; Score 35; DB 2; Length 88;  
Best Local Similarity 53.3%; Pred. No. 2.36+02;  
Matches 8; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 3 LMAMNIIISKEKEIK 17  
||:||||:||||  
Db 48 LMKMGFIKENDYK 62

RESULT 14

AB2227  
hypothetical protein asr3369 [imported] - Nostoc sp. (strain PCC 7120)  
C/Species: Nostoc sp. PCC 7120

A/Note: Nostoc sp. strain PCC 7120 is a synonym of Anabaena sp. strain PCC 7120  
C/Date: 14-Dec-2001 #sequence\_revision 14-Dec-2001 #text\_change 09-Dec-2002  
C/Accession: AB2227  
R/Kaneko, T.; Nakamura, Y.; Wolk, C.P.; Kuritz, T.; Sasamoto, S.; Watanabe, A.; Iriguchi,  
Nakazaki, N.; Shampo, S.; Sugimoto, M.; Takazawa, M.; Yamada, M.; Yasuda, M.; Tabata, S.  
DNA Res. 8, 205-213, 2001  
A/Title: Complete Genomic Sequence of the Filamentous Nitrogen-fixing Cyanobacterium Anat  
A/Reference number: AB1807; MUID:21595285; PMID:11759840  
A/Accession: AB2227  
A/Status: preliminary  
A/Molecule type: DNA  
A/Residues: 1-60 <KUR>  
A/Cross-references: GB:BA000019; PIDN:BAB75068.1; PID:gl7132464; GSPDB:GN00179  
A/Experimental source: strain PCC 7120  
C/Genetics:  
A/Gene: asr3369

Query Match 34.2%; Score 34.5; DB 2; Length 60;  
Best Local Similarity 44.4%; Pred. No. 1.96+02;  
Matches 8; Conservative 5; Mismatches 4; Indels 1; Gaps 1;

QY 2 VLMAMNIIIS-KEKKEIKW 18  
||:||||:||||  
Db 17 ILIAPNIEAIKERRNQKW 34

RESULT 15

A33058  
B32b delta-like toxin - Staphylococcus epidermidis  
C/Species: Staphylococcus epidermidis  
C/Date: 21-Sep-1990 #sequence\_revision 21-Sep-1990 #text\_change 19-Jan-1996  
C/Accession: A33058; A37189  
R/McKevitt, A.I.; Bjornson, G.L.; Mauracher, C.A.; Scheifele, D.W.  
submitted to the Protein Sequence Database, January 1990  
A/Reference number: A33058  
A/Accession: A33058  
A/Status: preliminary  
A/Molecule type: protein  
A/Residues: 1-25 <MCK>  
R/McKevitt, A.I.; Bjornson, G.L.; Mauracher, C.A.; Scheifele, D.W.  
Infect. Immun. 58, 1473-1475, 1990  
A/Title: Amino acid sequence of a delta-like toxin from Staphylococcus epidermidis.  
A/Reference number: A37189; MUID:90216022; PMID:2323825  
A/Accession: A37189  
A/Status: preliminary  
A/Molecule type: protein  
A/Residues: 1-5,7-25 <MC2>  
C/Superfamily: delta-hemolysin

Query Match 33.7%; Score 34; DB 2; Length 25;  
Best Local Similarity 50.0%; Pred. No. 90;  
Matches 8; Conservative 2; Mismatches 6; Indels 0; Gaps 0;

QY 4 MAMNIIISKEKEIKWI 19  
||:||||:||||  
Db 1 MAADIISTIGDLVKWI 16

RESULT 16

T45629  
hypothetical protein F13G24.240 - Arabidopsis thaliana  
C/Species: Arabidopsis thaliana (mouse-ear cress)  
C/Date: 04-Feb-2000 #sequence\_revision 04-Feb-2000 #text\_change 04-Feb-2000  
C/Accession: T45629  
R/Bevan, M.; Van Der Schueren, J.; Chuang, Y.J.; Voet, M.; Robben, J.; Volckaert, G.; Bar  
submitted to the Protein Sequence Database, December 1999  
A/Reference number: 223009  
A/Accession: T45629  
A/Status: preliminary  
A/Molecule type: DNA  
A/Residues: 1-54 <BEV>  
A/Cross-references: EMBL:AL133421  
A/Experimental source: cultivar Columbia; BAC clone F13G24





A;Title: Sequence analysis of the genome of Bombyx mori nucleopolyhedrovirus.

A;Reference number: Z22020; MUID:99281911; PMID:10355780

A;Accession: T41855

A;Status: preliminary; translated from GB/EMBL/DBJ

A;Molecule type: DNA

A;Residues: 1-95 <KAM>

A;Cross-references: EMBL:L33180; PIDN:AAC63784.1

A;Experimental source: isolate T3

C;Genetics:

A;Note: Orf\_96

Query Match 33.7%; Score 34; DB 2; Length 95;  
Best Local Similarity 36.8%; Pred. No. 3.6e+02;  
Matches 7; Conservative 7; Mismatches 3; Indels 2; Gaps 1;

QY 1 NVLMAMNIISKEKKEIKWI 19

DB 6 NVLLVPNAL--KRDVKYI 22

RESULT 27

B84142

hypothetical protein BH3938 [imported] - Bacillus halodurans (strain C-125)

C;Species: Bacillus halodurans

C;Date: 01-Dec-2000 #sequence\_revision 01-Dec-2000 #text\_change 15-Jun-2001

C;Accession: B84142

R;Takami, H.; Nakasone, K.; Takaki, Y.; Maeno, G.; Sasaki, R.; Masui, N.; Fujii, F.; Hira

Nucleic Acids Res. 28, 4317-4331, 2000

A;Title: Complete genome sequence of the alkaliphilic bacterium Bacillus halodurans and

A;Reference number: A83650; MUID:20512582; PMID:11058132

A;Accession: B84142

A;Status: preliminary

A;Molecule type: DNA

A;Residues: 1-96 <STO>

A;Cross-references: GB:AP001520; GB:BA000004; NID:gl0176401; PIDN:BAB07657.1; GSPDB:GN00

A;Experimental source: strain C-125

C;Genetics:

A;Gene: BH3938

Query Match 33.7%; Score 34; DB 2; Length 96;  
Best Local Similarity 42.1%; Pred. No. 3.7e+02;  
Matches 8; Conservative 5; Mismatches 6; Indels 0; Gaps 0;

QY 2 VLMAMNIISKEKKEIKWIG 20

DB 14 IVVLLNNIISKEKKPSSMG 32

RESULT 28

C45681

hypothetical 11K protein (gene 61.3 protein) - phage T4

N;Alternate names: gp 61.3

C;Species: phage T4

A;Note: host Escherichia coli

C;Date: 21-Sep-1993 #sequence\_revision 18-Nov-1994 #text\_change 08-Oct-1999

C;Accession: C45681; JS0560

R;Salick, H.E.; Stormo, G.D.; Dyson, R.L.; Alberts, B.M.

J. Virol. 67, 2305-2316, 1993

A;Title: Analysis of five presumptive protein-coding sequences clustered between the pri

A;Reference number: A45681; MUID:93186183; PMID:8383243

A;Accession: C45681

A;Status: preliminary

A;Molecule type: DNA

A;Residues: 1-97 <SEL>

A;Cross-references: GB:S57514; NID:g298518; PIDN:AB25710.1; PID:g298521

A;Note: sequence extracted from NCBI backbone (NCBIN:128289, NCBIP:128292)

C;Genetics:

A;Gene: 61.3

A;Map position: 19.765-20.056

Query Match 33.7%; Score 34; DB 2; Length 97;  
Best Local Similarity 41.7%; Pred. No. 3.7e+02;  
Matches 5; Conservative 5; Mismatches 2; Indels 0; Gaps 0;

QY 6 MNIIISKEKKEIK 17

DB 49 MNIVKDRPEMK 60

RESULT 29

H71893

probable molybdopterin converting factor, chain 1 - Helicobacter pylori (strain J99)

C;Species: Helicobacter pylori

A;Variety: strain J99

C;Date: 12-Feb-1999 #sequence\_revision 12-Feb-1999 #text\_change 19-Jan-2001

C;Accession: H71893

R;Alm, R.A.; Ling, L.S.L.; Moir, D.T.; King, B.L.; Brown, E.D.; Doig, P.C.; Smith, D.R.; J

Ives, C.; Gibson, R.; Merberg, D.; Mills, S.D.; Jiang, Q.; Taylor, D.E.; Vovis, G.F.; J

Nature 397, 176-180, 1999

A;Title: Genomic sequence comparison of two unrelated isolates of the human gastric patho

A;Reference number: A71800; MUID:99120557; PMID:9923682

A;Accession: H71893

A;Status: preliminary

A;Molecule type: DNA

A;Residues: 1-73 <ARN>

A;Cross-references: GB:AE001505; GB:AE001439; NID:g4155295; PIDN:AAD06323.1; PID:g415530

A;Experimental source: strain J99

C;Genetics:

A;Gene: moad

C;Superfamily: probable sulfur carrier protein Cj1517

Query Match 33.2%; Score 33.5; DB 2; Length 73;  
Best Local Similarity 38.1%; Pred. No. 3.3e+02;  
Matches 8; Conservative 5; Mismatches 7; Indels 1; Gaps 1;

QY 1 NVLMAMNIISKEKKEIK-WIG 20

DB 20 NGLKELRAILOEKEGLKEWLG 40

RESULT 30

A64620

molybdopterin converting factor, subunit 1 - Helicobacter pylori (strain 26695)

C;Species: Helicobacter pylori

C;Date: 09-Aug-1997 #sequence\_revision 09-Aug-1997 #text\_change 19-Jan-2001

C;Accession: A64620

R;Tomb, J.F.; White, O.; Kerlavage, A.R.; Clayton, R.A.; Sutton, G.G.; Fleischmann, R.D.;

Peterson, S.; Loftus, B.; Richardson, D.; Dodson, R.; Khalak, H.G.; Glodek, A.; McKenne

son, J.D.; Kelley, J.M.; Cotton, M.D.; Weidman, J.M.; Fujii, C.; Bowman, C.; Watthey, L.

Nature 388, 539-547, 1997

A;Authors: Wallin, E.; Hayes, W.S.; Borodovsky, M.; Karpk, P.D.; Smith, H.O.; Fraser, C.

A;Title: The complete genome sequence of the gastric pathogen Helicobacter pylori.

A;Reference number: A64520; MUID:97394467; PMID:9252185

A;Accession: A64620

A;Status: preliminary; nucleic acid sequence not shown; translation not shown

A;Molecule type: DNA

A;Residues: 1-74 <TON>

A;Cross-references: GB:AE000592; GB:AE000511; NID:g2313929; PIDN:AAD07857.1; PID:g231393

C;Superfamily: probable sulfur carrier protein Cj1517

P;74/Modified site: 1-thioglycine (Gly) #status predicted

Query Match 33.2%; Score 33.5; DB 2; Length 74;  
Best Local Similarity 38.1%; Pred. No. 3.4e+02;  
Matches 8; Conservative 5; Mismatches 7; Indels 1; Gaps 1;

QY 1 NVLMAMNIISKEKKEIK-WIG 20

DB 21 NDLKELRAILOEKEGLKEWLG 41

RESULT 31

E84146

hypothetical protein BH3973 [imported] - Bacillus halodurans (strain C-125)

C;Species: Bacillus halodurans

C;Date: 01-Dec-2000 #sequence\_revision 01-Dec-2000 #text\_change 15-Jun-2001

C;Accession: E84146

QY 1 NVLM--AMNII SKEKE-----IKWI 19

A;Cross-references: EMBL:AF020713; NID:g3025478; PID:g3025574; PIDN:AAC13069.1  
 R;Kunster, F.; Ogasawara, N.; Moszer, I.; Albertini, A.M.; Alloni, G.; Azevedo, V.; Bercet  
 C.; Bron, S.; Brulliet, S.; Bruchli, C.V.; Caldwell, B.; Capuano, V.; Carter, N.M.; Chd  
 A.; Ehrlich, S.D.; Emmerson, P.T.; Entian, K.D.; Errington, J.; Fabbret, C.; Ferrari, E.  
 Nature 390, 249-256, 1997  
 A;Authors: Foulger, D.; Fritz, C.; Fujita, M.; Fujita, Y.; Fuma, S.; Galizzi, A.; Gallen  
 iech, J.; Harwood, C.R.; Henaut, A.; Hilbert, H.; Holsappel, S.; Hosono, S.; Hullo, M.F.  
 Koetter, P.; Koningsstein, G.; Krogh, S.; Kumano, M.; Kurita, K.; Lapidus, A.; Lardinois,  
 A;Authors: Lauber, J.; Lazarevic, V.; Lee, S.M.; Levine, A.; Liu, H.; Masuda, S.; Maueel  
 Y, M.; Ogawa, K.; Ogiwara, A.; Oudega, B.; Park, S.H.; Parro, V.; Pohl, T.M.; Portetelle  
 Rieger, M.; Rivolta, C.; Rocha, E.; Roche, B.; Rose, M.; Sadaie, Y.; Sato, T.; Scanlon,  
 A;Authors: Schleich, S.; Schroeter, R.; Scoffone, P.; Sekiguchi, J.; Sekowska, A.; Seron  
 akuchin, M.; Tamakoshi, A.; Tanaka, T.; Terpstra, P.; Tognoni, A.; Tosato, V.; Uchiyama,  
 T.; Winters, P.; Wipat, A.; Yamamoto, H.; Yamane, K.; Yasumoto, K.; Yata, K.; Yoshida, K  
 A;Authors: Yoshikawa, H.F.; Zumstein, E.; Yoshikawa, H.; Danchin, A.  
 A;Title: The complete genome sequence of the Gram-positive bacterium *Bacillus subtilis*.  
 A;Reference number: A69580; MUID:98044033; PMID:9384377  
 A;Accession: H69318  
 A;Status: nucleic acid sequence not shown; translation not shown  
 A;Molecule type: DNA  
 A;Residues: 1-67 <KUN>  
 A;Cross-references: GB:Z99114; GB:Z99115; GB:AL009126; NID:g2634478; PIDN:CAB13989.1; PI  
 A;Experimental source: strain 168  
 C;Genetics:  
 A;Gene: yopZ

Query Match 32.7%; Score 33; DB 2; Length 67;  
 Best Local Similarity 41.7%; Pred. No. 3.6e+02;  
 Matches 5; Conservative 5; Mismatches 2; Indels 0; Gaps 0;  
 QY 6 MNIIISKEKKEIK 17  
 :::|||:::  
 Db 11 IDVIEKENKELR 22

RESULT 37  
 F97798  
 hypothetical protein RC0790 [imported] - Rickettsia conorii (strain Malish 7)  
 C;Species: Rickettsia conorii  
 C;Date: 30-Sep-2001 #sequence\_revision 30-Sep-2001 #text\_change 30-Sep-2001  
 C;Accession: F97798  
 R;Ogata, H.; Audic, S.; Renesto-Audiffren, P.; Fournier, P.E.; Barbe, V.; Samson, D.; Ri  
 Science 293, 2093-2098, 2001  
 A;Title: Mechanisms of Evolution in Rickettsia conorii and Rickettsia prowazekii.  
 A;Reference number: A97700; MUID:21442074; PMID:11557893  
 A;Accession: F97798  
 A;Status: preliminary  
 A;Molecule type: DNA  
 A;Residues: 1-67 <KUR>  
 A;Cross-references: GB:AE006914; PIDN:AAL03328.1; PID:g15619888; GSPDB:GN00173  
 C;Genetics:  
 A;Gene: RC0790

Query Match 32.7%; Score 33; DB 2; Length 67;  
 Best Local Similarity 41.7%; Pred. No. 3.6e+02;  
 Matches 5; Conservative 6; Mismatches 1; Indels 0; Gaps 0;  
 QY 7 NIISKEKKEIKW 18  
 :|||:::  
 Db 46 HIIDKEKEKURY 57

RESULT 38  
 H90086  
 hypothetical protein orf68 [imported] - Guillardia theta nucleomorph  
 C;Species: nucleomorph Guillardia theta  
 A;Note: a nucleomorph is the vestigial nucleus of a eukaryotic endosymbiont  
 C;Date: 10-May-2001 #sequence\_revision 10-May-2001 #text\_change 24-May-2001  
 C;Accession: H90086  
 R;Douglas, S.; Zauner, S.; Fraunholz, M.; Beaton, M.; Penny, S.; Deng, L.T.; Wu, X.; Rei  
 Nature 410, 1091-1096, 2001  
 A;Title: The highly reduced genome of an enlaved algal nucleus.  
 A;Reference number: A99082; MUID:11323671; PMID:11323671

A;Accession: H90086  
 A;Status: preliminary  
 A;Molecule type: DNA  
 A;Residues: 1-68 <DOU>  
 A;Cross-references: GB:AF165818; NID:g13794452; PIDN:AAK39827.1; GSPDB:GN00150  
 C;Genetics:  
 A;Gene: orf68  
 A;Map position: 1  
 A;Genome: nucleomorph  
 C;Keywords: nucleomorph

Query Match 32.7%; Score 33; DB 2; Length 68;  
 Best Local Similarity 50.0%; Pred. No. 3.7e+02;  
 Matches 7; Conservative 3; Mismatches 4; Indels 0; Gaps 0;  
 QY 4 MAMNIISKEKKEIK 17  
 :|||:::  
 Db 1 MVMNSKNKNEKKIK 14

RESULT 39  
 B71673  
 cold shock-like protein (cspA) RP670 - Rickettsia prowazekii  
 C;Species: Rickettsia prowazekii  
 C;Date: 21-Nov-1998 #sequence\_revision 21-Nov-1998 #text\_change 03-Nov-2000  
 C;Accession: B71673  
 R;Andersson, S.G.E.; Zomorodipour, A.; Andersson, J.O.; Sicheritz-Ponten, T.; Alsmark, U  
 Nature 396, 133-140, 1998  
 A;Title: The genome sequence of Rickettsia prowazekii and the origin of mitochondria.  
 A;Reference number: A71630; MUID:99039499; PMID:9823893  
 A;Accession: B71673  
 A;Status: preliminary; nucleic acid sequence not shown; translation not shown  
 A;Molecule type: DNA  
 A;Residues: 1-70 <AND>  
 A;Cross-references: GB:AJ235272; GB:AJ235269; NID:g3861033; PIDN:CAA15108.1; PID:g3861208  
 A;Experimental source: strain Madrid E  
 C;Genetics:  
 A;Gene: cspA; RP670  
 C;Superfamily: major cold shock protein; cold shock domain homology  
 F;7-67/Domain: cold shock domain homology <CSD>

Query Match 32.7%; Score 33; DB 2; Length 70;  
 Best Local Similarity 46.7%; Pred. No. 3.8e+02;  
 Matches 7; Conservative 2; Mismatches 2; Indels 4; Gaps 1;  
 QY 4 MAMNIISKEKKEIKW 18  
 :|||:::  
 Db 1 MATNIVGK---VKW 11

RESULT 40  
 E97827  
 cold shock-like protein [imported] - Rickettsia conorii (strain Malish 7)  
 C;Species: Rickettsia conorii  
 C;Date: 30-Sep-2001 #sequence\_revision 30-Sep-2001 #text\_change 17-May-2002  
 C;Accession: E97827  
 R;Ogata, H.; Audic, S.; Renesto-Audiffren, P.; Fournier, P.E.; Barbe, V.; Samson, D.; Ro  
 Science 293, 2093-2098, 2001  
 A;Title: Mechanisms of Evolution in Rickettsia conorii and Rickettsia prowazekii.  
 A;Reference number: A97700; MUID:21442074; PMID:11557893  
 A;Accession: E97827  
 A;Status: preliminary  
 A;Molecule type: DNA  
 A;Residues: 1-70 <KUR>  
 A;Cross-references: GB:AE006914; PIDN:AAL03559.1; PID:g15620137; GSPDB:GN00173  
 C;Genetics:  
 A;Gene: cspA  
 C;Superfamily: major cold shock protein; cold shock domain homology

Query Match 32.7%; Score 33; DB 2; Length 70;  
 Best Local Similarity 46.7%; Pred. No. 3.8e+02;  
 Matches 7; Conservative 2; Mismatches 2; Indels 4; Gaps 1;

QY 4 MAMNIISKEKEIKW 18  
|||:|:|:  
Db 1 MATNIVGK---VKW 11

RESULT 41  
JQ1805  
B1LR protein - vaccinia virus (strain WR)  
C:Species: vaccinia virus  
C>Date: 30-Sep-1993 #sequence\_revision 30-Sep-1993 #text\_change 08-Oct-1999  
C:Accession: JQ1805  
R:Smith, G.L.; Chan, Y.S.; Howard, S.T.  
J. Gen. Virol. 72, 1349-1376, 1991  
A:Title: Nucleotide sequence of 42kbp of vaccinia virus strain WR from near the right in  
A:Reference number: JQ1767; MUID:91259063; PMID:2045793  
A:Accession: JQ1805  
A:Molecule type: DNA  
A:Residues: 1-72 <SMI>  
A:CROSS-references: DDBJ:D11079; NID:G222717; PIDN:BA01841.1; PID:dl002317; PID:G222756

Query Match 32.7%; Score 33; DB 2; Length 72;  
Best Local Similarity 50.0%; Pred. No. 3.9e+02;  
Matches 8; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

QY 1 NVLMAMNIISKEKEIK 16  
|||:|:|:|:  
Db 7 NVEDIINEIDREKEEI 22

RESULT 42  
H90332  
hypothetical protein SS08760 [imported] - Sulfolobus solfataricus transposon ISCL491  
C:Species: Sulfolobus solfataricus  
C>Date: 24-May-2001 #sequence\_revision 24-May-2001 #text\_change 24-May-2001  
C:Accession: H90332  
R:She, Q.; Singh, R.K.; Confalonieri, F.; Zivanovic, Y.; Allard, G.; Awayez, M.J.; Chan-  
Jong, I.; Jeffries, A.C.; Kozera, C.J.; Medina, N.; Peng, X.; Thi-Ngoc, H.P.; Redder, R.  
arrett, R.A.; Ragan, M.A.; Senses, C.W.; Van der Oost, J.  
submitted to GenBank, April 2001  
A:Description: Sulfolobus solfataricus complete genome.  
A:Reference number: A99139  
A:Accession: H90332  
A>Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-76 <NUR>  
A:CROSS-references: GB:AE006641; NID:gl3814963; PIDN:AAK41919.1; GSPDB:GN00155  
C:Genetics:  
A:Gene: SS08760.

Query Match 32.7%; Score 33; DB 2; Length 76;  
Best Local Similarity 37.5%; Pred. No. 4.1e+02;  
Matches 6; Conservative 7; Mismatches 3; Indels 0; Gaps 0;

QY 2 VVKAELVITQEKREVE 36  
|||:|:|:|:  
Db 21 VVKAELVITQEKREVE 36

RESULT 43  
C90115  
small nuclear ribonucleoprotein SM D2 [imported] - Guillardia theta nucleomorph  
C:Species: nucleomorph Guillardia theta  
A:Note: a nucleomorph is the vestigial nucleus of a eukaryotic endosymbiont  
C>Date: 10-May-2001 #sequence\_revision 10-May-2001 #text\_change 24-May-2001  
C:Accession: C90115  
R:Douglas, S.; Zauner, S.; Fraunholz, M.; Beaton, M.; Penny, S.; Deng, L.T.; Wu, X.; Re-  
Nature 410, 1091-1096, 2001  
A:Title: The highly reduced genome of an enslaved algal nucleus.  
A:Reference number: A99082; MUID:11323671; PMID:11323671  
A:Accession: C90115  
A>Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-82 <DOU>

A:CROSS-references: GB:AJ010592; NID:g12580778; PIDN:CAC27096.1; GSPDB:GN00151  
C:Genetics:  
A:Map position: 2  
A:Genome: nucleomorph  
C:Keywords: nucleomorph

Query Match 32.7%; Score 33; DB 2; Length 82;  
Best Local Similarity 33.3%; Pred. No. 4.5e+02;  
Matches 5; Conservative 6; Mismatches 4; Indels 0; Gaps 0;

QY 6 MNIISKEKEIKWIG 20  
|||:|:|:|:  
Db 38 LNTLLENVNEIKFVG 52

RESULT 44  
F81321  
hypothetical protein Cj1164c [imported] - Campylobacter jejuni (strain NCTC 11168)  
C:Species: Campylobacter jejuni  
C>Date: 31-Mar-2000 #sequence\_revision 31-Mar-2000 #text\_change 03-Jun-2002  
C:Accession: F81321  
R:Parkhill, J.; Wren, B.W.; Mungall, K.; Ketley, J.M.; Churcher, C.; Bagham, D.; Chilling-  
C.W.; Quail, M.; Rajandream, M.A.; Rutherford, K.M.; VanVliet, A.; Whitehead, S.; Barrell  
Nature 403, 665-668, 2000  
A:Title: The genome sequence of the food-borne pathogen Campylobacter jejuni reveals hyp-  
A:Reference number: A81250; MUID:20150912; PMID:10698204  
A:Accession: F81321  
A>Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-87 <PAR>  
A:CROSS-references: GB:AL139077; GB:AL111168; NID:G5968444; PIDN:CAB73418.1; PID:G596859  
A:Experimental source: serotype O2, strain NCTC 11168  
C:Genetics:  
A:Gene: Cj1164c

Query Match 32.7%; Score 33; DB 2; Length 87;  
Best Local Similarity 50.0%; Pred. No. 4.7e+02;  
Matches 5; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 11 KEKKEIKWIG 20  
|||:|:|:|:  
Db 73 KKKKESWLG 82

RESULT 45  
A42527  
B1LR protein - vaccinia virus (strain Copenhagen)  
C:Species: vaccinia virus  
A:Note: host Homo sapiens (man)  
C>Date: 09-Nov-1990 #sequence\_revision 09-Nov-1990 #text\_change 08-Apr-1994  
C:Accession: A42527  
R:Johnson, G.P.  
submitted to GenBank, June 1990  
A:Reference number: A33172  
A:Accession: A42527  
A>Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-88 <JOH>

Query Match 32.7%; Score 33; DB 2; Length 88;  
Best Local Similarity 50.0%; Pred. No. 4.8e+02;  
Matches 8; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

QY 1 NVLMAMNIISKEKEIK 16  
|||:|:|:|:  
Db 23 NVEDIINEIDREKEEI 38

Search completed: February 17, 2004, 10:58:17  
Job time : 14.0792 secs



GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: February 17, 2004, 10:53:48 ; Search time 23.4455 Seconds  
(without alignments)  
142.889 Million cell updates/sec

Title: US-09-900-147-5

Perfect score: 76

Sequence: 1 RYVDALNVLMMNNIIS 16

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 801455 seqs, 209382283 residues

Total number of hits satisfying chosen parameters: 318354

Minimum DB seq length: 0

Maximum DB seq length: 100

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications AA:\*

1: /cgn2\_6/ptodata/1/pubpaa/US07\_PUBCOMB.pep.\*  
2: /cgn2\_6/ptodata/1/pubpaa/PCT\_NEW\_PUB.pep.\*  
3: /cgn2\_6/ptodata/1/pubpaa/US06\_NEW\_PUB.pep.\*  
4: /cgn2\_6/ptodata/1/pubpaa/US06\_PUBCOMB.pep.\*  
5: /cgn2\_6/ptodata/1/pubpaa/US07\_NEW\_PUB.pep.\*  
6: /cgn2\_6/ptodata/1/pubpaa/PCTUS\_PUBCOMB.pep.\*  
7: /cgn2\_6/ptodata/1/pubpaa/US08\_NEW\_PUB.pep.\*  
8: /cgn2\_6/ptodata/1/pubpaa/US08\_PUBCOMB.pep.\*  
9: /cgn2\_6/ptodata/1/pubpaa/US09A\_PUBCOMB.pep.\*  
10: /cgn2\_6/ptodata/1/pubpaa/US09B\_PUBCOMB.pep.\*  
11: /cgn2\_6/ptodata/1/pubpaa/US09C\_PUBCOMB.pep.\*  
12: /cgn2\_6/ptodata/1/pubpaa/US09\_NEW\_PUB.pep.\*  
13: /cgn2\_6/ptodata/1/pubpaa/US10A\_PUBCOMB.pep.\*  
14: /cgn2\_6/ptodata/1/pubpaa/US10B\_PUBCOMB.pep.\*  
15: /cgn2\_6/ptodata/1/pubpaa/US10C\_PUBCOMB.pep.\*  
16: /cgn2\_6/ptodata/1/pubpaa/US10\_NEW\_PUB.pep.\*  
17: /cgn2\_6/ptodata/1/pubpaa/US60\_NEW\_PUB.pep.\*  
18: /cgn2\_6/ptodata/1/pubpaa/US60\_PUBCOMB.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	76	100.0	16	10	US-09-900-147-5
2	76	100.0	19	10	US-09-900-147-3
3	76	100.0	37	10	US-09-900-147-1
4	76	100.0	74	15	US-10-214-188-10
5	68	89.5	19	10	US-09-900-147-15
6	68	89.5	19	10	US-09-900-147-16
7	67	88.2	30	10	US-09-900-147-6
8	64	84.2	14	10	US-09-900-147-11
9	64	84.2	19	10	US-09-900-147-17
10	46	60.5	20	10	US-09-900-147-4
11	46	60.5	96	9	US-09-864-761-45697
12	44	57.9	11	10	US-09-900-147-9
13	42	55.3	9	10	US-09-900-147-2
14	39	51.3	74	15	US-10-214-188-5
15	39	51.3	74	15	US-10-214-188-7

16	39	51.3	75	15	US-10-214-188-9	Sequence 9, Appli
17	39	51.3	76	12	US-10-029-386-29071	Sequence 29071, A
18	39	51.3	85	11	US-09-932-581-5	Sequence 5, Appli
19	39	51.3	85	12	US-10-338-294-5	Sequence 5, Appli
20	39	51.3	85	14	US-10-165-614-2	Sequence 2, Appli
21	38	50.0	69	15	US-10-214-188-8	Sequence 8, Appli
22	38	50.0	74	15	US-10-214-188-6	Sequence 6, Appli
23	38	50.0	76	11	US-09-932-581-6	Sequence 6, Appli
24	38	50.0	76	12	US-10-338-294-6	Sequence 6, Appli
25	38	50.0	76	14	US-10-165-614-3	Sequence 3, Appli
26	35	46.1	63	11	US-09-764-872-344	Sequence 344, App
27	35	46.1	90	9	US-09-925-299-824	Sequence 824, App
28	35	46.1	90	11	US-09-925-299-824	Sequence 824, App
29	34	44.7	59	15	US-10-036-542-148	Sequence 148, App
30	34	44.7	67	9	US-09-864-761-43263	Sequence 43263, A
31	34	44.7	80	9	US-09-801-574-48	Sequence 48, Appli
32	34	44.7	88	9	US-09-939-980-513	Sequence 513, App
33	33	43.4	41	9	US-09-864-761-40202	Sequence 40202, A
34	33	43.4	59	14	US-10-044-359-10	Sequence 10, Appli
35	33	43.4	73	9	US-09-864-761-41902	Sequence 41902, A
36	32	42.1	58	14	US-10-044-359-2	Sequence 2, Appli
37	32	42.1	97	12	US-09-933-767-812	Sequence 812, App
38	32	42.1	97	15	US-10-023-282-812	Sequence 812, App
39	31.5	41.4	99	9	US-09-864-761-46653	Sequence 46653, A
40	31	40.8	7	10	US-09-900-147-7	Sequence 7, Appli
41	31	40.8	15	11	US-09-966-459A-55	Sequence 55, Appli
42	31	40.8	15	11	US-09-966-422B-70	Sequence 70, Appli
43	31	40.8	15	12	US-10-262-272A-70	Sequence 70, Appli
44	31	40.8	19	12	US-09-962-756-1214	Sequence 1214, Ap
45	31	40.8	19	12	US-10-253-471-1214	Sequence 1214, Ap

#### ALIGNMENTS

RESULT 1  
US-09-900-147-5  
; Sequence 5, Application US/09900147  
; Patent No. US20020103121A1  
; GENERAL INFORMATION:  
; APPLICANT: La Thangue, Nicholas B  
; TITLE OF INVENTION: Peptide antagonists of DP transcription factors  
; FILE REFERENCE: 620-67  
; CURRENT FILING DATE: 2001-07-09  
; PRIOR FILING DATE: 2001-07-09  
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 09/308,935  
; PRIOR FILING DATE: EARLIER FILING DATE: 1999-05-27  
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: GB 9626589.7  
; NUMBER OF SEQ ID NOS: 18  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 5  
; LENGTH: 16  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Description of Artificial Sequence: Synthetic  
US-09-900-147-5

Query Match 100.0%; Score 76; DB 10; Length 16;  
Best Local Similarity 100.0%; Pred. No. 6.4e-07;  
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 RYVDALNVLMMNNIIS 16  
|||  
Db 1 RYVDALNVLMMNNIIS 16

RESULT 2  
US-09-900-147-3  
; Sequence 3, Application US/09900147  
; Patent No. US20020103121A1

Tue Feb 17 11:55:52 2004

```

; GENERAL INFORMATION:
; APPLICANT: La Thangue, Nicholas B
; APPLICANT: Bandara, Laseantha R
; TITLE OF INVENTION: Peptide antagonists of DP transcription factors
; FILE REFERENCE: 620-67
; CURRENT APPLICATION NUMBER: US/09/900,147
; CURRENT FILING DATE: 2001-07-09
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 09/308,935
; PRIOR FILING DATE: EARLIER FILING DATE: 1999-05-27
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: GB 9626589.7
; PRIOR FILING DATE: EARLIER FILING DATE: 1996-12-20
; NUMBER OF SEQ ID NOS: 18
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 3
; LENGTH: 19
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Synthetic
; US-09-900-147-3

Query Match 100.0%; Score 76; DB 10; Length 19;
Best Local Similarity 100.0%; Pred. No. 7.8e-07;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 RYVDALNVLMMNNIIS 16
Db 3 RYVDALNVLMMNNIIS 18

RESULT 3
US-09-900-147-1
; Sequence 1, Application US/09900147
; Patent No. US20020103121A1
; GENERAL INFORMATION:
; APPLICANT: La Thangue, Nicholas B
; APPLICANT: Bandara, Laseantha R
; TITLE OF INVENTION: Peptide antagonists of DP transcription factors
; FILE REFERENCE: 620-67
; CURRENT APPLICATION NUMBER: US/09/900,147
; CURRENT FILING DATE: 2001-07-09
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 09/308,935
; PRIOR FILING DATE: EARLIER FILING DATE: 1999-05-27
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: GB 9626589.7
; PRIOR FILING DATE: EARLIER FILING DATE: 1996-12-20
; NUMBER OF SEQ ID NOS: 18
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1
; LENGTH: 37
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Synthetic
; US-09-900-147-1

Query Match 100.0%; Score 76; DB 10; Length 37;
Best Local Similarity 100.0%; Pred. No. 1.7e-06;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 RYVDALNVLMMNNIIS 16
Db 6 RYVDALNVLMMNNIIS 21

RESULT 4
US-10-214-188-10
; Sequence 10, Application US/10214188
; Publication No. US2003002260A1
; GENERAL INFORMATION:
; APPLICANT: LA THANGUE, NICHOLAS B.
; APPLICANT: BERNARDS, RENE
; APPLICANT: HUMANS, ELEANORE M.
; TITLE OF INVENTION: TRANSCRIPTION FACTOR E2F-5
; US-09-900-147-5
```

```

; NUMBER OF SEQUENCES: 25
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: NIXON & VANDERHVE P.C.
; STREET: 1100 NORTH GLEBE ROAD
; CITY: ARLINGTON
; STATE: VIRGINIA
; COUNTRY: U.S.A.
; ZIP: 22201-4714
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/10/214,188
; FILING DATE: 08-Aug-2002
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/894,139
; FILING DATE: 13-AUG-1997
; ATTORNEY/AGENT INFORMATION:
; NAME: WILSON, MARY J.
; REGISTRATION NUMBER: 32,955
; REFERENCE/DOCKET NUMBER: 620-22
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703) 816-4000
; TELEFAX: (703) 816-4100
; INFORMATION FOR SEQ ID NO: 10:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 74 amino acids
; TYPE: amino acid
; STRANDEDNESS: <Unknown>
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; SEQUENCE DESCRIPTION: SEQ ID NO: 10:
; US-10-214-188-10

Query Match 100.0%; Score 76; DB 15; Length 74;
Best Local Similarity 100.0%; Pred. No. 3.7e-06;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 RYVDALNVLMMNNIIS 16
Db 48 RYVDALNVLMMNNIIS 63

RESULT 5
US-09-900-147-15
; Sequence 15, Application US/09900147
; Patent No. US20020103121A1
; GENERAL INFORMATION:
; APPLICANT: La Thangue, Nicholas B
; APPLICANT: Bandara, Laseantha R
; TITLE OF INVENTION: Peptide antagonists of DP transcription factors
; FILE REFERENCE: 620-67
; CURRENT APPLICATION NUMBER: US/09/900,147
; CURRENT FILING DATE: 2001-07-09
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 09/308,935
; PRIOR FILING DATE: EARLIER FILING DATE: 1999-05-27
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: GB 9626589.7
; PRIOR FILING DATE: EARLIER FILING DATE: 1996-12-20
; NUMBER OF SEQ ID NOS: 18
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 15
; LENGTH: 19
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Mutant peptide
; US-09-900-147-15

Query Match 89.5%; Score 68; DB 10; Length 19;
Best Local Similarity 87.5%; Pred. No. 2e-05;
```

Matches 14; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 RYDNLNVLMMNNIIS 16  
| | | | | | | | | |  
Db 3 RYDNLNVLMMNNIIS 18

## RESULT 6

US-09-900-147-16

; Sequence 16, Application US/09900147  
; Patent No. US20020103121A1  
; GENERAL INFORMATION:  
; APPLICANT: La Thangue, Nicholas B  
; TITLE OF INVENTION: Peptide antagonists of DP transcription factors  
; FILE REFERENCE: 620-67  
; CURRENT APPLICATION NUMBER: US/09/900,147  
; CURRENT FILING DATE: 2001-07-09  
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 09/308,935  
; PRIOR FILING DATE: EARLIER FILING DATE: 1999-05-27  
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: GB 9626589.7  
; PRIOR FILING DATE: EARLIER FILING DATE: 1996-12-20  
; NUMBER OF SEQ ID NOS: 18  
; SOFTWARE: Patentin Ver. 2.1  
; SEQ ID NO 16  
; TYPE: PRT  
; LENGTH: 19  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Description of Artificial Sequence: Mutant peptide  
US-09-900-147-16

Query Match 89.5%; Score 68; DB 10; Length 19;  
Best Local Similarity 93.8%; Pred. No. 2e-05;  
Matches 15; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 RYDNLNVLMMNNIIS 16  
| | | | | | | | | |  
Db 3 RYDNLNVLMMNNIIS 18

## RESULT 7

US-09-900-147-6

; Sequence 6, Application US/09900147  
; Patent No. US20020103121A1  
; GENERAL INFORMATION:  
; APPLICANT: La Thangue, Nicholas B  
; TITLE OF INVENTION: Peptide antagonists of DP transcription factors  
; FILE REFERENCE: 620-67  
; CURRENT APPLICATION NUMBER: US/09/900,147  
; CURRENT FILING DATE: 2001-07-09  
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 09/308,935  
; PRIOR FILING DATE: EARLIER FILING DATE: 1999-05-27  
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: GB 9626589.7  
; PRIOR FILING DATE: EARLIER FILING DATE: 1996-12-20  
; NUMBER OF SEQ ID NOS: 18  
; SOFTWARE: Patentin Ver. 2.1  
; SEQ ID NO 6  
; TYPE: PRT  
; LENGTH: 30  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Description of Artificial Sequence: Synthetic  
US-09-900-147-6

Query Match 88.2%; Score 67; DB 10; Length 30;  
Best Local Similarity 100.0%; Pred. No. 5.1e-05;  
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 YDNLNVLMMNNIIS 16  
| | | | | | | | | |  
Db 1 YDNLNVLMMNNIIS 14

## RESULT 8

US-09-900-147-11

; Sequence 11, Application US/09900147  
; Patent No. US20020103121A1  
; GENERAL INFORMATION:  
; APPLICANT: La Thangue, Nicholas B  
; TITLE OF INVENTION: Peptide antagonists of DP transcription factors  
; FILE REFERENCE: 620-67  
; CURRENT APPLICATION NUMBER: US/09/900,147  
; CURRENT FILING DATE: 2001-07-09  
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 09/308,935  
; PRIOR FILING DATE: EARLIER FILING DATE: 1999-05-27  
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: GB 9626589.7  
; PRIOR FILING DATE: EARLIER FILING DATE: 1996-12-20  
; NUMBER OF SEQ ID NOS: 18  
; SOFTWARE: Patentin Ver. 2.1  
; SEQ ID NO 11  
; TYPE: PRT  
; LENGTH: 14  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Description of Artificial Sequence: Synthetic  
US-09-900-147-11

Query Match 84.2%; Score 64; DB 10; Length 14;  
Best Local Similarity 100.0%; Pred. No. 7.2e-05;  
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RYDNLNVLMMNNIIS 13  
| | | | | | | | | |  
Db 2 RYDNLNVLMMNNIIS 14

## RESULT 9

US-09-900-147-17

; Sequence 17, Application US/09900147  
; Patent No. US20020103121A1  
; GENERAL INFORMATION:  
; APPLICANT: La Thangue, Nicholas B  
; TITLE OF INVENTION: Peptide antagonists of DP transcription factors  
; FILE REFERENCE: 620-67  
; CURRENT APPLICATION NUMBER: US/09/900,147  
; CURRENT FILING DATE: 2001-07-09  
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 09/308,935  
; PRIOR FILING DATE: EARLIER FILING DATE: 1999-05-27  
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: GB 9626589.7  
; PRIOR FILING DATE: EARLIER FILING DATE: 1996-12-20  
; NUMBER OF SEQ ID NOS: 18  
; SOFTWARE: Patentin Ver. 2.1  
; SEQ ID NO 17  
; TYPE: PRT  
; LENGTH: 19  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Description of Artificial Sequence: Mutant peptide  
US-09-900-147-17

Query Match 84.2%; Score 64; DB 10; Length 19;  
Best Local Similarity 87.5%; Pred. No. 0.0001;  
Matches 14; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 RYDNLNVLMMNNIIS 16  
| | | | | | | | | |  
Db 3 RYDNLNVLMMNNIIS 18

## RESULT 10

US-09-900-147-4

; Sequence 4, Application US/09900147

```

; Patent No. US20020103121A1
;
; GENERAL INFORMATION:
;
; APPLICANT: La Thangue, Nicholas B
; APPLICANT: Bandara, Lasantha R
;
; TITLE OF INVENTION: Peptide antagonists of DP transcription factors
;
; FILE REFERENCE: 620-67
;
; CURRENT APPLICATION NUMBER: US/09/900,147
;
; CURRENT FILING DATE: 2001-07-09
;
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 09/308,935
;
; PRIOR FILING DATE: EARLIER FILING DATE: 1999-05-27
;
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: GB 9626589.7
;
; PRIOR FILING DATE: EARLIER FILING DATE: 1996-12-20
;
; NUMBER OF SEQ ID NOS: 18
;
; SOFTWARE: PatentIn Ver. 2.1
;
; SEQ ID NO 4
;
; LENGTH: 20
;
; TYPE: PRT
;
; ORGANISM: Artificial Sequence
;
; FEATURE:
;
; OTHER INFORMATION: Description of Artificial Sequence: Synthetic
;
; US-09-900-147-4

```

Query Match	60.5%	Score 46;	DB 10;	Length 20;
Best Local Similarity	100.0%	Pred. No. 0.16;		
Matches	10;	Conservative	0;	Mismatches 0;
		Indels	0;	Gaps 0;

Qy	7 NVLMAMNIIS 16
Db	1 NVLMAMNIIS 10

```

RESULT 11
US/09-864-761-45697
; Sequence 45697, Application US/09864761
; Patent No. US20020048763A1
; GENERAL INFORMATION:
; APPLICANT: Penn, Sharon G.
; APPLICANT: Rank, David R.
; APPLICANT: Hanzel, David K.
; APPLICANT: Chen, Wensheng
; TITLE OF INVENTION: HUMAN GENOME-DERIVED
; TITLE OF INVENTION: GENE EXPRESSION ANAL
; FILE REFERENCE: Aomics-X-1
; CURRENT APPLICATION NUMBER: US/09/864,761
; CURRENT FILING DATE: 2001-05-23
; PRIOR APPLICATION NUMBER: US 60/180,312
; PRIOR FILING DATE: 2000-02-04
; PRIOR APPLICATION NUMBER: US 60/207,456
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: US 09/632,366
; PRIOR FILING DATE: 2000-08-03
; PRIOR APPLICATION NUMBER: GB 24263.6
; PRIOR FILING DATE: 2000-10-04
; PRIOR APPLICATION NUMBER: US 60/236,359
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: PCT/US01/00666
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00667
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00664
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00669
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00665
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00668
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00663
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00662
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00661
; PRIOR FILING DATE: 2001-01-30

```

```

, PRIOR APPLICATION NUMBER: PCT/US01/00670
, PRIOR FILING DATE: 2001-01-30
, PRIOR APPLICATION NUMBER: US 60/234,687
, PRIOR FILING DATE: 2000-09-21
, PRIOR APPLICATION NUMBER: US 09/608,408
, PRIOR FILING DATE: 2000-06-30
, PRIOR APPLICATION NUMBER: US 09/774,203
, PRIOR FILING DATE: 2001-01-29
, NUMBER OF SEQ ID NOS: 49117
, SOFTWARE: Annonax Sequence Listing Engine vers. 1.1
, SEQ ID NO 45697
, LENGTH: 96
, TYPE: PRT
, ORGANISM: Homo sapiens
, FEATURE:
, OTHER INFORMATION: MAP TO AC021804.3
, OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 0.89
, OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 0.99
, OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 0.9
, OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 0.92
, OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 0.88
, OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 0.75
, OTHER INFORMATION: EST HUMAN HIT: BB880658.1, EVALUAE 2.00e-51
, OTHER INFORMATION: SWISSPROT HIT: Q61501, EVALUAE 5.00e-06
, OS-09-864-761-45697

```

Query Match 60.5%; Score 46; DB 9; Length 96;  
Best Local Similarity 43.8%; Pred. No. 0.98;  
Matches 7; Conservative 7; Mismatches 2; Indels

Qy 1 RYDALNVLMMNIIS 16  
 . | : | | : | | : : : |  
pb 5 RYDITVNLESLHVS 20

```

RESULT 12
US-09-900-147-9
; Sequence 9, Application US/09900147
; Patent No. US20020103121A1
; GENERAL INFORMATION:
; APPLICANT: La Thangue, Nicholas B
; APPLICANT: Bandaru, Lasantha R
; TITLE OF INVENTION: Peptide antagonists of DP transduction factors
; FILE REFERENCE: 620-67
; CURRENT APPLICATION NUMBER: US/09/900,147
; CURRENT FILING DATE: 2001-07-09
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 09/308,935
; PRIOR FILING DATE: EARLIER FILING DATE: 1999-05-27
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: GB 9626589.7
; PRIOR FILING DATE: EARLIER FILING DATE: 1996-12-20
; NUMBER OF SEQ ID NOS: 18
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 9
; LENGTH: 11
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Synthetic
; US-09-900-147-9

```

Query Match 57.9%; Score 44; DB 10; Length 11;  
Best Local Similarity 100.0%; Pred. No. 0.19;  
Matches 9: Conservative 0: Mismatches 0: Indels 0: Gaps 0;

Qy 1 RVDALNVL 9  
|||  
Db: 3 RVDALNVL 11

RESULT 13  
US-09-900-147-2  
; Sequence 2, Application US/09900147  
; Patent No. US20020103121A1

```
; GENERAL INFORMATION:
; APPLICANT: La Thangue, Nicholas B
; APPLICANT: Bandara, Lasantha R
; TITLE OF INVENTION: Peptide antagonists of DP transcription factors
; FILE REFERENCE: 620-67
; CURRENT APPLICATION NUMBER: US/09/900,147
; CURRENT FILING DATE: 2001-07-09
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 09/308,935
; PRIOR FILING DATE: EARLIER FILING DATE: 1999-05-27
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: GB 962589.7
; PRIOR FILING DATE: EARLIER FILING DATE: 1996-12-20
; NUMBER OF SEQ ID NOS: 18
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2
; LENGTH: 9
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Synthetic
US-09-900-147-2

Query Match 55.3%; Score 42; DB 10; Length 9;
Best Local Similarity 100.0%; Pred. No. 7e+05;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 7 NVLMANNII 15
Db 1 NVLMANNII 9

RESULT 14
US-10-214-188-5
; Sequence 5, Application US/10214188
; Publication No. US20030022260A1
; GENERAL INFORMATION:
; APPLICANT: LA THANGUE, NICHOLAS B.
; BERNARDS, RENE
; HIJMAN, ELEANORE M.
; TITLE OF INVENTION: TRANSCRIPTION FACTOR E2F-5
; NUMBER OF SEQUENCES: 25
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: NIXON & VANDERHYE P.C.
; STREET: 1100 NORTH GLEBE ROAD
; CITY: ARLINGTON
; STATE: VIRGINIA
; COUNTRY: U.S.A.
; ZIP: 22201-4714
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/10/214,188
; FILING DATE: 08-Aug-2002
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/894,139
; FILING DATE: 13-AUG-1997
; ATTORNEY/AGENT INFORMATION:
; NAME: WILSON, MARY J.
; REGISTRATION NUMBER: 32,955
; REFERENCE/DOCKET NUMBER: 620-22
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703) 816-4000
; TELEFAX: (703) 816-4100
; INFORMATION FOR SEQ ID NO: 5:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 74 amino acids
; TYPE: amino acid
; STRANDEDNESS: <Unknown>
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; SEQUENCE DESCRIPTION: SEQ ID NO: 5:
US-10-214-188-7

Query Match 55.3%; Score 39; DB 15; Length 74;
Best Local Similarity 43.8%; Pred. No. 13;
Matches 7; Conservative 4; Mismatches 5; Indels 0; Gaps 0;

QY 1 RYVDALNVLMANNII 16
Db 47 RYDITNVLEGIQLIA 62

RESULT 15
US-10-214-188-7
; Sequence 7, Application US/10214188
; Publication No. US20030022260A1
; GENERAL INFORMATION:
; APPLICANT: LA THANGUE, NICHOLAS B.
; BERNARDS, RENE
; HIJMAN, ELEANORE M.
; TITLE OF INVENTION: TRANSCRIPTION FACTOR E2F-5
; NUMBER OF SEQUENCES: 25
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: NIXON & VANDERHYE P.C.
; STREET: 1100 NORTH GLEBE ROAD
; CITY: ARLINGTON
; STATE: VIRGINIA
; COUNTRY: U.S.A.
; ZIP: 22201-4714
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/10/214,188
; FILING DATE: 08-Aug-2002
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/894,139
; FILING DATE: 13-AUG-1997
; ATTORNEY/AGENT INFORMATION:
; NAME: WILSON, MARY J.
; REGISTRATION NUMBER: 32,955
; REFERENCE/DOCKET NUMBER: 620-22
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703) 816-4000
; TELEFAX: (703) 816-4100
; INFORMATION FOR SEQ ID NO: 7:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 74 amino acids
; TYPE: amino acid
; STRANDEDNESS: <Unknown>
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; SEQUENCE DESCRIPTION: SEQ ID NO: 7:
US-10-214-188-9

Query Match 51.3%; Score 39; DB 15; Length 74;
Best Local Similarity 46.7%; Pred. No. 13;
Matches 7; Conservative 4; Mismatches 4; Indels 0; Gaps 0;

QY 1 RYVDALNVLMANNII 15
Db 47 RYDITNVLEGIHLI 61

RESULT 16
US-10-214-188-9
; Sequence 9, Application US/10214188
; Publication No. US20030022260A1
; GENERAL INFORMATION:
; APPLICANT: LA THANGUE, NICHOLAS B.
```

```
; SEQUENCE DESCRIPTION: SEQ ID NO: 5:
US-10-214-188-5

Query Match 51.3%; Score 39; DB 15; Length 74;
Best Local Similarity 43.8%; Pred. No. 13;
Matches 7; Conservative 4; Mismatches 5; Indels 0; Gaps 0;

QY 1 RYVDALNVLMANNII 16
Db 47 RYDITNVLEGIQLIA 62

RESULT 15
US-10-214-188-7
; Sequence 7, Application US/10214188
; Publication No. US20030022260A1
; GENERAL INFORMATION:
; APPLICANT: LA THANGUE, NICHOLAS B.
; BERNARDS, RENE
; HIJMAN, ELEANORE M.
; TITLE OF INVENTION: TRANSCRIPTION FACTOR E2F-5
; NUMBER OF SEQUENCES: 25
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: NIXON & VANDERHYE P.C.
; STREET: 1100 NORTH GLEBE ROAD
; CITY: ARLINGTON
; STATE: VIRGINIA
; COUNTRY: U.S.A.
; ZIP: 22201-4714
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/10/214,188
; FILING DATE: 08-Aug-2002
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/894,139
; FILING DATE: 13-AUG-1997
; ATTORNEY/AGENT INFORMATION:
; NAME: WILSON, MARY J.
; REGISTRATION NUMBER: 32,955
; REFERENCE/DOCKET NUMBER: 620-22
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703) 816-4000
; TELEFAX: (703) 816-4100
; INFORMATION FOR SEQ ID NO: 7:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 74 amino acids
; TYPE: amino acid
; STRANDEDNESS: <Unknown>
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; SEQUENCE DESCRIPTION: SEQ ID NO: 7:
US-10-214-188-7

Query Match 51.3%; Score 39; DB 15; Length 74;
Best Local Similarity 46.7%; Pred. No. 13;
Matches 7; Conservative 4; Mismatches 4; Indels 0; Gaps 0;

QY 1 RYVDALNVLMANNII 15
Db 47 RYDITNVLEGIHLI 61

RESULT 16
US-10-214-188-9
; Sequence 9, Application US/10214188
; Publication No. US20030022260A1
; GENERAL INFORMATION:
; APPLICANT: LA THANGUE, NICHOLAS B.
```

```

; BERNARDS, RENE
; HIJWANS, ELEANORE M.
; TITLE OF INVENTION: TRANSCRIPTION FACTOR E2F-5
; NUMBER OF SEQUENCES: 25
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: NIXON & VANDERHIE P.C.
; STREET: 1100 NORTH GLEBE ROAD
; CITY: ARLINGTON
; STATE: VIRGINIA
; COUNTRY: U.S.A.
; ZIP: 22201-4714
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/10/214,188
; FILING DATE: 08-Aug-2002
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/894,139
; FILING DATE: 13-AUG-1997
; ATTORNEY/AGENT INFORMATION:
; NAME: WILSON, MARY J.
; REGISTRATION NUMBER: 32,955
; REFERENCE/DOCKET NUMBER: 620-22
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703) 816-4000
; TELEFAX: (703) 816-4100
; INFORMATION FOR SEQ ID NO: 9:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 75 amino acids
; TYPE: amino acid
; STRANDEDNESS: <Unknown>
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; SEQUENCE DESCRIPTION: SEQ ID NO: 9:
US-10-214-188-9

Query Match 51.3%; Score 39; DB 15; Length 75;
Best Local Similarity 46.7%; Pred. No. 13;
Matches 7; Conservative 4; Mismatches 4; Indels 0; Gaps 0;

Qy 1 RYDANLVLMAMNII 15
Db 48 RIYDITNVLEGIDLI 62

RESULT 17
US-10-029-386-29071
; Sequence 29071, Application US/10029386
; Publication No. US20030194704A1
; GENERAL INFORMATION:
; APPLICANT: Penn, Sharon G.
; APPLICANT: Rank, David K.
; APPLICANT: Hanzel, David K.
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR G
; FILE REFERENCE: AEOMICA-X-2
; CURRENT APPLICATION NUMBER: US/10/029,386
; CURRENT FILING DATE: 2001-12-20
; NUMBER OF SEQ ID NOS: 34288
; SOFTWARE: Anomax Sequence Listing Engine vers. 1.1
; SEQ ID NO 29071
; LENGTH: 76
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: MAP TO CHR20.1
; OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 6.1
; OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 5.4
; OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 4.6

```

```

; OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 3.3
; OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 4.8
; OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 4.4
; OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 4
; OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 5.4
; OTHER INFORMATION: SWISSPROT HIT: Q01094, EVALUE 1.00e-36
US-10-029-386-29071

Query Match 51.3%; Score 39; DB 12; Length 76;
Best Local Similarity 43.8%; Pred. No. 13;
Matches 7; Conservative 4; Mismatches 5; Indels 0; Gaps 0;

Qy 1 RYDANLVLMAMNII 16
Db 51 RIYDITNVLEGIGLIA 66

RESULT 18
US-09-932-581-5
; Sequence 5, Application US/09932581
; Publication No. US20030050264A1
; GENERAL INFORMATION:
; APPLICANT: Andrews, William H.
; APPLICANT: Foster, Christopher A.
; APPLICANT: Fraser, Stephanie
; APPLICANT: Mohammadpour, Hamid
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR MODULATING
; TITLE OF INVENTION: TELOMERASE REVERSE TRANSCRIPTASE (TERT) EXPRESSION
; FILE REFERENCE: SIER-005
; CURRENT APPLICATION NUMBER: US/09/932,581
; CURRENT FILING DATE: 2001-08-17
; PRIOR APPLICATION NUMBER: 60/227,865
; PRIOR FILING DATE: 2000-08-24
; PRIOR APPLICATION NUMBER: 60/230,174
; PRIOR FILING DATE: 2000-09-01
; PRIOR APPLICATION NUMBER: 60/238,345
; PRIOR FILING DATE: 2000-10-05
; NUMBER OF SEQ ID NOS: 25
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 5
; LENGTH: 85
; TYPE: PRT
; ORGANISM: human
US-09-932-581-5

Query Match 51.3%; Score 39; DB 11; Length 85;
Best Local Similarity 43.8%; Pred. No. 15;
Matches 7; Conservative 4; Mismatches 5; Indels 0; Gaps 0;

Qy 1 RYDANLVLMAMNII 16
Db 57 RIYDITNVLEGIGLIA 72

RESULT 19
US-10-338-294-5
; Sequence 5, Application US/10338294
; Publication No. US20030171326A1
; GENERAL INFORMATION:
; APPLICANT: Andrews, William H.
; APPLICANT: Foster, Christopher A.
; APPLICANT: Fraser, Stephanie
; APPLICANT: Mohammadpour, Hamid
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR MODULATING
; TITLE OF INVENTION: TELOMERASE REVERSE TRANSCRIPTASE (TERT) EXPRESSION
; FILE REFERENCE: SIER-005
; CURRENT APPLICATION NUMBER: US/10/338,294
; CURRENT FILING DATE: 2003-01-07
; PRIOR APPLICATION NUMBER: US/09/932,581
; PRIOR FILING DATE: 2001-08-17
; PRIOR APPLICATION NUMBER: 60/227,865
; PRIOR FILING DATE: 2000-08-24
; PRIOR APPLICATION NUMBER: 60/230,174

```

```
; PRIOR FILING DATE: 2000-09-01
; PRIOR APPLICATION NUMBER: 60/238,345
; PRIOR FILING DATE: 2000-10-05
; NUMBER OF SEQ ID NOS: 25
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 5
; LENGTH: 85
; TYPE: PRT
; ORGANISM: human
US-10-338-294-5

Query Match      51.3%; Score 39; DB 12; Length 85;
Best Local Similarity 43.8%; Pred. No. 15;
Matches 7; Conservative 4; Mismatches 5; Indels 0; Gaps 0;

QY 1 RYVDALNVLMMNII 16
Db 57 RYDITNVLEGIQLIA 72

RESULT 20
US-10-165-614-2
; Sequence 2, Application US/10165614
; Publication No. US20020193289A1
; GENERAL INFORMATION:
; APPLICANT: Andrews, William H.
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR MODULATING
; FILE OF INVENTION: TELOMERASE REVERSE TRANSCRIPTASE (TERT) EXPRESSION
; FILE REFERENCE: SIER-018
; CURRENT APPLICATION NUMBER: US/10/165,614
; CURRENT FILING DATE: 2002-06-25
; PRIOR APPLICATION NUMBER: 60/296,992
; PRIOR FILING DATE: 2001-06-07
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 85
; TYPE: PRT
; ORGANISM: homo sapiens
US-10-165-614-2

Query Match      51.3%; Score 39; DB 14; Length 85;
Best Local Similarity 43.8%; Pred. No. 15;
Matches 7; Conservative 4; Mismatches 5; Indels 0; Gaps 0;

QY 1 RYVDALNVLMMNII 16
Db 57 RYDITNVLEGIQLIA 72

RESULT 21
US-10-214-188-8
; Sequence 8, Application US/10214188
; Publication No. US20030022260A1
; GENERAL INFORMATION:
; APPLICANT: LA THANGUE, NICHOLAS B.
; HILJMAN, ELEANORE M.
; TITLE OF INVENTION: TRANSCRIPTION FACTOR E2F-5
; NUMBER OF SEQUENCES: 25
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: NIXON & VANDERHUYE P.C.
; STREET: 1100 NORTH GLEBE ROAD
; CITY: ARLINGTON
; STATE: VIRGINIA
; COUNTRY: U.S.A.
; ZIP: 22201-4714
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/10/214,188
; FILING DATE: 08-Aug-2002
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/894,139
; FILING DATE: 13-AUG-1997
; ATTORNEY/AGENT INFORMATION:
; NAME: WILSON, MARY J.
; REGISTRATION NUMBER: 32,955
; REFERENCE/DOCKET NUMBER: 620-22
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703) 816-4000
; TELEFAX: (703) 816-4100
; INFORMATION FOR SEQ ID NO: 6:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 74 amino acids
; TYPE: amino acid
; STRANDEDNESS: <Unknown>
```

Tue Feb 17 11:55:52 2004

```
;
;
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; SEQUENCE DESCRIPTION: SEQ ID NO: 6:
US-10-214-188-6

Query Match      50.0%; Score 38; DB 15; Length 74;
Best Local Similarity 46.7%; Pred. No. 19;
Matches 7; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

QY 1 RYVDALNVLMAMNII 15
Db 47 RIYDITNVLEGIGLI 61

RESULT 23
US-09-932-581-6
; Sequence 6, Application US/09932581
; Publication No. US20030050264A1
; GENERAL INFORMATION:
; APPLICANT: Andrews, William H.
; APPLICANT: Foster, Christopher A.
; APPLICANT: Fraser, Stephanie
; APPLICANT: Mohammadpour, Hamid
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR MODULATING
; TITLE OF INVENTION: TELOMERASE REVERSE TRANSCRIPTASE (TERT) EXPRESSION
; FILE REFERENCE: SIER-005
; CURRENT APPLICATION NUMBER: US/09/932,581
; CURRENT FILING DATE: 2001-08-17
; PRIOR APPLICATION NUMBER: 60/227,865
; PRIOR FILING DATE: 2000-08-24
; PRIOR APPLICATION NUMBER: 60/230,174
; PRIOR FILING DATE: 2000-09-01
; PRIOR APPLICATION NUMBER: 60/238,345
; PRIOR FILING DATE: 2000-10-05
; NUMBER OF SEQ ID NOS: 25
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 6
; LENGTH: 76
; TYPE: PRT
; ORGANISM: human
US-09-932-581-6

Query Match      50.0%; Score 38; DB 11; Length 76;
Best Local Similarity 46.7%; Pred. No. 20;
Matches 7; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

QY 1 RYVDALNVLMAMNII 15
Db 47 RIYDITNVLEGIGLI 61

RESULT 24
US-10-338-294-6
; Sequence 6, Application US/10338294
; Publication No. US20030171326A1
; GENERAL INFORMATION:
; APPLICANT: Andrews, William H.
; APPLICANT: Foster, Christopher A.
; APPLICANT: Fraser, Stephanie
; APPLICANT: Mohammadpour, Hamid
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR MODULATING
; TITLE OF INVENTION: TELOMERASE REVERSE TRANSCRIPTASE (TERT) EXPRESSION
; FILE REFERENCE: SIER-005
; CURRENT APPLICATION NUMBER: US/10/338,294
; CURRENT FILING DATE: 2003-01-07
; PRIOR APPLICATION NUMBER: US/09/932,581
; PRIOR FILING DATE: 2001-08-17
; PRIOR APPLICATION NUMBER: 60/227,865
; PRIOR FILING DATE: 2000-08-24
; PRIOR APPLICATION NUMBER: 60/230,174
; PRIOR FILING DATE: 2000-09-01
; PRIOR APPLICATION NUMBER: 60/238,345
; PRIOR FILING DATE: 2000-10-05
; NUMBER OF SEQ ID NOS: 25
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 6
; LENGTH: 76
; TYPE: PRT
; ORGANISM: human
US-10-338-294-6

Query Match      50.0%; Score 38; DB 12; Length 76;
Best Local Similarity 46.7%; Pred. No. 20;
Matches 7; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

QY 1 RYVDALNVLMAMNII 15
Db 47 RIYDITNVLEGIGLI 61

RESULT 25
US-10-165-614-3
; Sequence 3, Application US/10165614
; Publication No. US20020193289A1
; GENERAL INFORMATION:
; APPLICANT: Andrews, William H.
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR MODULATING
; TITLE OF INVENTION: TELOMERASE REVERSE TRANSCRIPTASE (TERT) EXPRESSION
; FILE REFERENCE: SIER-018
; CURRENT APPLICATION NUMBER: US/10/165,614
; CURRENT FILING DATE: 2002-06-25
; PRIOR APPLICATION NUMBER: 60/296,992
; PRIOR FILING DATE: 2001-06-07
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 3
; LENGTH: 76
; TYPE: PRT
; ORGANISM: homo sapiens
US-10-165-614-3

Query Match      50.0%; Score 38; DB 14; Length 76;
Best Local Similarity 46.7%; Pred. No. 20;
Matches 7; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

QY 1 RYVDALNVLMAMNII 15
Db 47 RIYDITNVLEGIGLI 61

RESULT 26
US-09-764-872-344
; Sequence 344, Application US/09764872
; Publication No. US20030050231A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PA125
; CURRENT APPLICATION NUMBER: US/09/764,872
; CURRENT FILING DATE: 2001-01-17
; Prior application data removed - consult PALM or file wrapper
; NUMBER OF SEQ ID NOS: 957
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 344
; LENGTH: 63
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (3)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: SITE
; LOCATION: (29)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; US-09-764-872-344
```



Query Match 46.1%; Score 35; DB 11; Length 63;  
Best Local Similarity 57.1%; Pred. No. 53;  
Matches 8; Conservative 1; Mismatches 5; Indels 0; Gaps 0;

Qy 2 VYDALNVLMMNII 15  
|| || || || : ||  
Db 8 VYELLNFMMLKNLI 21

**RESULT 27**

US-09-925-299-824  
; Sequence 824, Application US/09925299  
; Patent No. US20020055627A1  
; GENERAL INFORMATION:  
; APPLICANT: Rosen et al.  
; TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies  
; FILE REFERENCE: PA102  
; CURRENT APPLICATION NUMBER: US/09/925,299  
; PRIOR FILING DATE: 2001-08-10  
; PRIOR APPLICATION NUMBER: PCT/US00/05883  
; PRIOR FILING DATE: 2000-03-08  
; PRIOR APPLICATION NUMBER: 60/124,270  
; PRIOR FILING DATE: 1999-03-12  
; NUMBER OF SEQ ID NOS: 1556  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 824  
; LENGTH: 90  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: SITE  
; LOCATION: (36)  
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids  
US-09-925-299-824

RESULT 28

```

US-09-925-299-824
; Sequence 824, Application US/09925299
; Publication No. US20030040617A9
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies
; FILE REFERENCE: PA102
; CURRENT APPLICATION NUMBER: US/09/925,299
; CURRENT FILING DATE: 2001-08-10
; PRIOR APPLICATION NUMBER: PCT/US00/05883
; PRIOR FILING DATE: 2000-03-08
; PRIOR APPLICATION NUMBER: 60/124,270
; PRIOR FILING DATE: 1999-03-12
; NUMBER OF SEQ ID NOS: 1556
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 824
; LENGTH: 90
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (36)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
US-09-925-299-824

```

Qy 3 YDALNVLM 10  
Dbb 42 FDALNVLM 49  
:||||:|

RESULT 29

```

US-10-036-542-148
; Sequence 148, Application US/10036542
; Publication No. US20030083481A1
; GENERAL INFORMATION:
; APPLICANT: Birse et al.
; TITLE OF INVENTION: 25 Human Prostate and Prostate Cancer Associated Proteins
; FILE REFERENCE: PA002PI
; CURRENT APPLICATION NUMBER: US/10/036,542
; CURRENT FILING DATE: 2002-01-07
; PRIOR APPLICATION NUMBER: PCT/US00/19666
; PRIOR FILING DATE: 2000-07-20
; PRIOR APPLICATION NUMBER: 60/144,972
; PRIOR FILING DATE: 1999-07-21
; PRIOR APPLICATION NUMBER: 60/148,681
; PRIOR FILING DATE: 1999-08-13
; PRIOR APPLICATION NUMBER: 60/149,173
; PRIOR FILING DATE: 1999-08-17
; PRIOR APPLICATION NUMBER: 60/158,004
; PRIOR FILING DATE: 1999-10-06
; PRIOR APPLICATION NUMBER: 60/194,689
; PRIOR FILING DATE: 2000-04-05
; NUMBER OF SEQ ID NOS: 157
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 148
; LENGTH: 59
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-036-542-148

```

## RESULT 30

US-09-864-761-43263  
Sequence 43263, Application US/09864761  
Patent NO. US20020048763A1  
GENERAL INFORMATION:  
APPLICANT: Penn, Sharron G.  
APPLICANT: Rank, David R.  
APPLICANT: Hanzel, David K.  
APPLICANT: Chen, Wensheng  
TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR  
TITLE OF INVENTION: GENE EXPRESSION ANALYSIS BY MICROARRAY  
FILE REFERENCE: Aemica-X-1  
CURRENT APPLICATION NUMBER: US/09/864,761  
CURRENT FILING DATE: 2001-05-23  
PRIOR APPLICATION NUMBER: US 60/180,312  
PRIOR FILING DATE: 2000-02-04  
PRIOR APPLICATION NUMBER: US 60/207,456  
PRIOR FILING DATE: 2000-05-26  
PRIOR APPLICATION NUMBER: US 09/532,366  
PRIOR FILING DATE: 2000-08-03  
PRIOR APPLICATION NUMBER: GB 24263.6  
PRIOR FILING DATE: 2000-10-04  
PRIOR APPLICATION NUMBER: US 60/236,359  
PRIOR FILING DATE: 2000-09-27  
PRIOR APPLICATION NUMBER: PCT/US01/00666  
PRIOR FILING DATE: 2001-01-30  
PRIOR APPLICATION NUMBER: PCT/US01/00667  
PRIOR FILING DATE: 2001-01-30

```

; PRIOR APPLICATION NUMBER: PCT/US01/00664
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00669
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00665
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00668
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00663
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00662
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00661
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00670
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: US 60/234,687
; PRIOR FILING DATE: 2000-09-21
; PRIOR APPLICATION NUMBER: US 09/608,408
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: US 09/774,203
; PRIOR FILING DATE: 2001-01-29
; NUMBER OF SEQ ID NOS: 49117
; SOFTWARE: Annomax Sequence Listing Engine vers. 1.1
; SEQ ID NO 43263
; LENGTH: 67
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: MAP TO AL031301.1
; OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 1.6
US-09-864-761-43263

Query Match 44.7%; Score 34; DB 9; Length 67;
Best Local Similarity 35.7%; Pred. No. 86;
Matches 5; Conservative 6; Mismatches 3; Indels 0; Gaps 0;

Qy 2 VYDALNVLMANNII 15
Db 38 IITTTIIIIANNII 51

RESULT 31
US-09-801-574-48
; Sequence 48, Application US/09801574
; Patent No. US20020081592A1
; GENERAL INFORMATION:
; APPLICANT: Wang, Peijiang Jeremy
; TITLE OF INVENTION: Reproduction-Specific Genes
; FILE REFERENCE: 0399.2007.002
; CURRENT APPLICATION NUMBER: US/09/801,574
; CURRENT FILING DATE: 2001-03-07
; PRIOR APPLICATION NUMBER: 60/187,518
; PRIOR FILING DATE: 2000-03-07
; PRIOR APPLICATION NUMBER: 60/261,557
; PRIOR FILING DATE: 2001-01-12
; NUMBER OF SEQ ID NOS: 90
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 48
; LENGTH: 80
; TYPE: PRT
; ORGANISM: Mus musculus
US-09-801-574-48

Query Match 44.7%; Score 34; DB 9; Length 80;
Best Local Similarity 50.0%; Pred. No. 1.1e+02;
Matches 7; Conservative 1; Mismatches 6; Indels 0; Gaps 0;

Qy 1 RYVDALNVLMANNI 14
Db 31 KYISVLNVAFACNI 44

RESULT 32
US-09-939-980-513
; Sequence 513, Application US/09939980
; Patent No. US2002008234A1
; GENERAL INFORMATION:
; APPLICANT: Black, Michael
; Burnham, Martin
; Hodgson, John
; Knowles, David
; Lonetto, Michael
; Nicholas, Richard
; Pratt, Julie
; Reichard, Richard
; Rosenberg, Martin
; Ward, Judith
; TITLE OF INVENTION: No. US2002008234A1el Prokaryotic Polynucleotides,
; Polypeptides and Their Uses
; NUMBER OF SEQUENCES: 534
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: SmithKline Beecham Corporation
; STREET: 709 Swedeland Road
; CITY: King of Prussia
; STATE: PA
; COUNTRY: USA
; ZIP: 19406-0939
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/939,980
; FILING DATE: 27-Aug-2001
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/936,165
; FILING DATE: <Unknown>
; ATTORNEY/AGENT INFORMATION:
; NAME: Gimmi, Edward R
; REGISTRATION NUMBER: 38,891
; REFERENCE/DOCKET NUMBER: P50549
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 610-270-4478
; TELEFAX: 610-270-5090
; TELEX: <Unknown>
; INFORMATION FOR SEQ ID NO: 513:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 88 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: Protein
; SEQUENCE DESCRIPTION: SEQ ID NO: 513:
US-09-939-980-513

Query Match 44.7%; Score 34; DB 9; Length 88;
Best Local Similarity 56.2%; Pred. No. 1.2e+02;
Matches 9; Conservative 3; Mismatches 2; Indels 2; Gaps 1;

Qy 2 VYDAL--NVLMANNII 15
Db 48 MYDAKMDNVLPINII 63

RESULT 33
US-09-864-761-40202
; Sequence 40202, Application US/09864761
; Patent No. US20020048763A1
; GENERAL INFORMATION:
; APPLICANT: Penn, Sharron G.
; APPLICANT: Rank, David R.
; APPLICANT: Hanzel, David K.
```

APPLICANT: Chen, Wensheng  
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR  
; FILE REFERENCE: Aomic-X-1  
; CURRENT APPLICATION NUMBER: US/09/864,761  
; PRIOR FILING DATE: 2001-05-23  
; PRIOR FILING DATE: 2000-02-04  
; PRIOR APPLICATION NUMBER: US 60/180,312  
; PRIOR FILING DATE: 2000-05-26  
; PRIOR APPLICATION NUMBER: US 09/632,366  
; PRIOR FILING DATE: 2000-08-03  
; PRIOR APPLICATION NUMBER: GB 24263.6  
; PRIOR FILING DATE: 2000-10-04  
; PRIOR APPLICATION NUMBER: US 60/236,359  
; PRIOR FILING DATE: 2000-09-27  
; PRIOR APPLICATION NUMBER: PCT/US01/00666  
; PRIOR FILING DATE: 2001-01-30  
; PRIOR APPLICATION NUMBER: PCT/US01/00667  
; PRIOR FILING DATE: 2001-01-30  
; PRIOR APPLICATION NUMBER: PCT/US01/00664  
; PRIOR FILING DATE: 2001-01-30  
; PRIOR APPLICATION NUMBER: PCT/US01/00669  
; PRIOR FILING DATE: 2001-01-30  
; PRIOR APPLICATION NUMBER: PCT/US01/00665  
; PRIOR FILING DATE: 2001-01-30  
; PRIOR APPLICATION NUMBER: PCT/US01/00668  
; PRIOR FILING DATE: 2001-01-30  
; PRIOR APPLICATION NUMBER: PCT/US01/00663  
; PRIOR FILING DATE: 2001-01-30  
; PRIOR APPLICATION NUMBER: PCT/US01/00662  
; PRIOR FILING DATE: 2001-01-30  
; PRIOR APPLICATION NUMBER: PCT/US01/00661  
; PRIOR FILING DATE: 2001-01-30  
; PRIOR APPLICATION NUMBER: PCT/US01/00670  
; PRIOR FILING DATE: 2001-01-30  
; PRIOR APPLICATION NUMBER: US 60/234,687  
; PRIOR FILING DATE: 2000-09-21  
; PRIOR APPLICATION NUMBER: US 09/608,408  
; PRIOR FILING DATE: 2000-06-30  
; PRIOR APPLICATION NUMBER: US 09/774,203  
; PRIOR FILING DATE: 2001-01-29  
; NUMBER OF SEQ ID NOS: 49117  
; SOFTWARE: Annomax Sequence Listing Engine vers. 1.1  
; SEQ ID NO 40202  
; LENGTH: 41  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
; FEATURE:  
; OTHER INFORMATION: MAP TO AC021000.3  
; OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 1.5  
; OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 1.4  
; OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 1.7  
; OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 5.2  
; OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 1.8  
; OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 1.8  
; OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 2.6  
; OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 1.8  
; OTHER INFORMATION: EST HUMAN HIT: AV724222.1, EVALUATE 9.00e-17  
; OTHER INFORMATION: SWISSPROT HIT: P54938, EVALUATE 7.00e-18  
US-09-864-761-40202

Query Match 43.4%; Score 33; DB 9; Length 41;  
Best Local Similarity 63.6%; Pred. No. 74;  
Matches 7; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 4 DALNVLMAMNI 14  
| | | | |  
Db 26 DLNVLMALDI 36

RESULT 34  
US-10-044-359-10

Sequence 10, Application US/10044359  
; Publication No. US20020160454A1  
; GENERAL INFORMATION:  
; APPLICANT: Herzman, Rafael  
; APPLICANT: Wong, James F.  
; APPLICANT: Lee, Jian-Ming  
; TITLE OF INVENTION: SCORPION TOXINS  
; FILE REFERENCE: BB1367 US NA  
; CURRENT APPLICATION NUMBER: US/10/044,359  
; CURRENT FILING DATE: 2002-01-11  
; PRIOR APPLICATION NUMBER: 09/599,416  
; PRIOR FILING DATE: 2000-06-22  
; PRIOR APPLICATION NUMBER: 60/140,227  
; PRIOR FILING DATE: 1999-06-22  
; NUMBER OF SEQ ID NOS: 28  
; SOFTWARE: Microsoft Office 97  
; SEQ ID NO 10  
; LENGTH: 59  
; TYPE: PRT  
; ORGANISM: Hottentotta judaica  
US-10-044-359-10  
Query Match 43.4%; Score 33; DB 14; Length 59;  
Best Local Similarity 33.3%; Pred. No. 1.1e+02;  
Matches 5; Conservative 6; Mismatches 4; Indels 0; Gaps 0;  
QY 1 RVYDALNVLMAMNII 15  
| | | | |  
Db 5 RLVAIILVLMNVI 19  
RESULT 35  
US-09-864-761-41902  
; Sequence 41902, Application US/09864761  
; Patent No. US20020048763A1  
; GENERAL INFORMATION:  
; APPLICANT: Penn, Sharron G.  
; APPLICANT: Rank, David R.  
; APPLICANT: Hanzel, David K.  
; APPLICANT: Chen, Wensheng  
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR  
; FILE REFERENCE: Aomic-X-1  
; CURRENT APPLICATION NUMBER: US/09/864,761  
; CURRENT FILING DATE: 2001-05-23  
; PRIOR APPLICATION NUMBER: US 60/180,312  
; PRIOR FILING DATE: 2000-02-04  
; PRIOR APPLICATION NUMBER: US 60/207,456  
; PRIOR FILING DATE: 2000-05-26  
; PRIOR APPLICATION NUMBER: US 09/632,366  
; PRIOR FILING DATE: 2000-08-03  
; PRIOR APPLICATION NUMBER: GB 24263.6  
; PRIOR FILING DATE: 2000-10-04  
; PRIOR APPLICATION NUMBER: US 60/236,359  
; PRIOR FILING DATE: 2000-09-27  
; PRIOR APPLICATION NUMBER: PCT/US01/00666  
; PRIOR FILING DATE: 2001-01-30  
; PRIOR APPLICATION NUMBER: PCT/US01/00667  
; PRIOR FILING DATE: 2001-01-30  
; PRIOR APPLICATION NUMBER: PCT/US01/00664  
; PRIOR FILING DATE: 2001-01-30  
; PRIOR APPLICATION NUMBER: PCT/US01/00669  
; PRIOR FILING DATE: 2001-01-30  
; PRIOR APPLICATION NUMBER: PCT/US01/00665  
; PRIOR FILING DATE: 2001-01-30  
; PRIOR APPLICATION NUMBER: PCT/US01/00668  
; PRIOR FILING DATE: 2001-01-30  
; PRIOR APPLICATION NUMBER: PCT/US01/00663  
; PRIOR FILING DATE: 2001-01-30  
; PRIOR APPLICATION NUMBER: PCT/US01/00662  
; PRIOR FILING DATE: 2001-01-30  
; PRIOR APPLICATION NUMBER: PCT/US01/00661  
; PRIOR FILING DATE: 2001-01-30

Tue Feb 17 11:55:52 2004

```

; PRIOR APPLICATION NUMBER: PCT/US01/00670
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: US 60/234,687
; PRIOR FILING DATE: 2000-09-21
; PRIOR APPLICATION NUMBER: US 09/608,408
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: US 09/774,203
; PRIOR FILING DATE: 2001-01-29
; NUMBER OF SEQ ID NOS: 49117
; SOFTWARE: Anomax Sequence Listing Engine vers. 1.1
; SEQ ID NO 41902
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: MAP TO AL133415.12
; OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 3.9
; OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 4.7
; OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 6.2
; OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 5.6
; OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 7.9
; OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 7.7
; OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 5.8
; OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 5.8
; OTHER INFORMATION: EST HUMAN HIT: AUL19666.1, EVALUE 6.00e-35
; OTHER INFORMATION: SWISSPROT HIT: P08670, EVALUE 6.00e-36
; US-09-864-761-41902

Query Match 43.4%; Score 33; DB 9; Length 73;
Best Local Similarity 63.6%; Pred. No. 1.4e+02;
Matches 7; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 4 DALNVLMAMNI 14
| | | | | | | | | | | | | | | | | | | | | |
Db 49 DLLNVKVALDI 59

RESULT 36
US-10-044-359-2
; Sequence 2, Application US/10044359
; Publication No. US20020160454A1
; GENERAL INFORMATION:
; APPLICANT: Hertman, Rafael
; APPLICANT: Wong, James F.
; APPLICANT: Lee, Jian-Ming
; TITLE OF INVENTION: SCORPION TOXINS
; FILE REFERENCE: B01367 US NA
; CURRENT APPLICATION NUMBER: US/10/044,359
; CURRENT FILING DATE: 2002-01-11
; PRIOR APPLICATION NUMBER: 09/599,416
; PRIOR FILING DATE: 2000-06-22
; PRIOR APPLICATION NUMBER: 60/140,227
; PRIOR FILING DATE: 1999-06-22
; NUMBER OF SEQ ID NOS: 28
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 2
; LENGTH: 58
; TYPE: PRT
; ORGANISM: Hottentotta judaica
US-10-044-359-2

Query Match 42.1%; Score 32; DB 14; Length 58;
Best Local Similarity 33.3%; Pred. No. 1.6e+02;
Matches 5; Conservative 6; Mismatches 4; Indels 0; Gaps 0;

QY 1 RYVDALNVLMAMNII 15
| : : : : | : | |
Db 3 RIITILIVFALNII 17

RESULT 37
US-09-933-767-812
; Sequence 812, Application US/09933767
; Publication No. US20030181692A1
; GENERAL INFORMATION:
; APPLICANT: Ni et al.
; TITLE OF INVENTION: 207 Human Secreted Proteins
; FILE REFERENCE: P2007P2
; CURRENT APPLICATION NUMBER: US/09/933,767
; CURRENT FILING DATE: 2001-08-22
; PRIOR APPLICATION NUMBER: PCT/US01/05614
; PRIOR FILING DATE: 2001-02-21
; PRIOR APPLICATION NUMBER: 60/184,836
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: 60/193,170
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: 09/205,258
; PRIOR FILING DATE: 1998-12-04
; PRIOR APPLICATION NUMBER: PCT/US98/11422
; PRIOR FILING DATE: 1998-06-04
; PRIOR APPLICATION NUMBER: 60/048,885
; PRIOR FILING DATE: 1997-06-06
; PRIOR APPLICATION NUMBER: 60/049,375
; PRIOR FILING DATE: 1997-06-06
; PRIOR APPLICATION NUMBER: 60/048,881
; PRIOR FILING DATE: 1997-06-06
; PRIOR APPLICATION NUMBER: 60/048,880
; PRIOR FILING DATE: 1997-06-06
; PRIOR APPLICATION NUMBER: 60/048,896
; PRIOR FILING DATE: 1997-06-06
; PRIOR APPLICATION NUMBER: 60/049,020
; PRIOR FILING DATE: 1997-06-06
; PRIOR APPLICATION NUMBER: 60/048,876
; PRIOR FILING DATE: 1997-06-06
; PRIOR APPLICATION NUMBER: 60/048,895
; PRIOR FILING DATE: 1997-06-06
; PRIOR APPLICATION NUMBER: 60/048,884
; PRIOR FILING DATE: 1997-06-06
; PRIOR APPLICATION NUMBER: 60/048,894
; PRIOR FILING DATE: 1997-06-06
; PRIOR APPLICATION NUMBER: 60/048,971
; PRIOR FILING DATE: 1997-06-06
; PRIOR APPLICATION NUMBER: 60/048,964
; PRIOR FILING DATE: 1997-06-06
; PRIOR APPLICATION NUMBER: 60/048,882
; PRIOR FILING DATE: 1997-06-06
; PRIOR APPLICATION NUMBER: 60/048,899
; PRIOR FILING DATE: 1997-06-06
; PRIOR APPLICATION NUMBER: 60/048,893
; PRIOR FILING DATE: 1997-06-06
; PRIOR APPLICATION NUMBER: 60/048,900
; PRIOR FILING DATE: 1997-06-06
; PRIOR APPLICATION NUMBER: 60/048,901
; PRIOR FILING DATE: 1997-06-06
; PRIOR APPLICATION NUMBER: 60/048,892
; PRIOR FILING DATE: 1997-06-06
; PRIOR APPLICATION NUMBER: 60/048,915
; PRIOR FILING DATE: 1997-06-06
; PRIOR APPLICATION NUMBER: 60/049,019
; PRIOR FILING DATE: 1997-06-06
; PRIOR APPLICATION NUMBER: 60/048,970
; PRIOR FILING DATE: 1997-06-06
; PRIOR APPLICATION NUMBER: 60/048,972
; PRIOR FILING DATE: 1997-06-06
; PRIOR APPLICATION NUMBER: 60/048,916
; PRIOR FILING DATE: 1997-06-06
; PRIOR APPLICATION NUMBER: 60/049,373
; PRIOR FILING DATE: 1997-06-06
; PRIOR APPLICATION NUMBER: 60/048,875
; PRIOR FILING DATE: 1997-06-06
; PRIOR APPLICATION NUMBER: 60/049,374
; PRIOR FILING DATE: 1997-06-06
; PRIOR APPLICATION NUMBER: 60/048,917
; PRIOR FILING DATE: 1997-06-06
; PRIOR APPLICATION NUMBER: 60/048,949
; PRIOR FILING DATE: 1997-06-06
```

; PRIOR APPLICATION NUMBER: 60/048,974
; PRIOR FILING DATE: 1997-06-06
; PRIOR APPLICATION NUMBER: 60/048,883
; PRIOR FILING DATE: 1997-06-06
; PRIOR APPLICATION NUMBER: 60/048,897
; PRIOR FILING DATE: 1997-06-06
; PRIOR APPLICATION NUMBER: 60/048,898
; PRIOR FILING DATE: 1997-06-06
; PRIOR APPLICATION NUMBER: 60/048,962
; PRIOR FILING DATE: 1997-06-06
; PRIOR APPLICATION NUMBER: 60/048,963
; PRIOR FILING DATE: 1997-06-06
; PRIOR APPLICATION NUMBER: 60/048,877
; PRIOR FILING DATE: 1997-06-06
; PRIOR APPLICATION NUMBER: 60/048,878
; PRIOR FILING DATE: 1997-06-06
; PRIOR APPLICATION NUMBER: 60/068,054
; PRIOR FILING DATE: 1997-12-18
; PRIOR APPLICATION NUMBER: 60/068,064
; PRIOR FILING DATE: 1997-12-18
; PRIOR APPLICATION NUMBER: 60/068,053
; PRIOR FILING DATE: 1997-12-18
; PRIOR APPLICATION NUMBER: 60/070,923
; PRIOR FILING DATE: 1997-12-18
; PRIOR APPLICATION NUMBER: 60/073,160
; PRIOR FILING DATE: 1998-01-30
; PRIOR APPLICATION NUMBER: 60/073,159
; PRIOR FILING DATE: 1998-01-30
; PRIOR APPLICATION NUMBER: 60/073,165
; PRIOR FILING DATE: 1998-01-30
; PRIOR APPLICATION NUMBER: 60/073,164
; PRIOR FILING DATE: 1998-01-30
; PRIOR APPLICATION NUMBER: 60/085,925
; PRIOR FILING DATE: 1998-05-18
; PRIOR APPLICATION NUMBER: 60/085,921
; PRIOR FILING DATE: 1998-05-18
; PRIOR APPLICATION NUMBER: 60/085,923
; PRIOR FILING DATE: 1998-05-18
; PRIOR APPLICATION NUMBER: 60/085,922
; PRIOR FILING DATE: 1998-05-18
; PRIOR APPLICATION NUMBER: 60/092,921
; PRIOR FILING DATE: 1998-07-15
; PRIOR APPLICATION NUMBER: 60/094,657
; PRIOR FILING DATE: 1998-07-30
; NUMBER OF SEQ ID NOS: 1245
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 812
; LENGTH: 97
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (16)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: SITE
; LOCATION: (38)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
US-09-933-767-812

Query Match 42.1%; Score 32; DB 12; Length 97;
Best Local Similarity 38.5%; Pred. No. 3e+02;
Matches 5; Conservative 5; Mismatches 3; Indels 0; Gaps 0;

QY 3 YDALNVLMMNII 15
Db 66 YDAIAVFLCIIHV 78

RESULT 38
US-10-023-282-812
; Sequence 812, Application US/10023282
; Publication No. US20030092893A1
; GENERAL INFORMATION:

; APPLICANT: Young et al.
; TITLE OF INVENTION: 207 Human Secreted Proteins
; FILE REFERENCE: P2007P1
; CURRENT APPLICATION NUMBER: US/10/023,282
; CURRENT FILING DATE: 2001-12-20
; EARLIER APPLICATION NUMBER: 09/205,258
; EARLIER FILING DATE: 1998-12-04
; EARLIER APPLICATION NUMBER: PCT/US98/11422
; EARLIER FILING DATE: 1998-06-04
; EARLIER APPLICATION NUMBER: 60/048,885
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/049,375
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,881
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,880
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,896
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/049,020
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,876
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,895
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,884
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,894
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,971
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,964
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,882
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,899
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,893
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,900
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,901
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,892
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,915
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/049,019
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,970
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,972
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,916
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/049,373
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,875
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/049,374
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,917
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,883
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,897
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,898
; EARLIER FILING DATE: 1997-06-06

```

; EARLIER APPLICATION NUMBER: 60/048,962
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,963
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,877
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,878
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/070,923
; EARLIER FILING DATE: 1997-12-18
; EARLIER APPLICATION NUMBER: 60/092,921
; EARLIER FILING DATE: 1998-07-15
; EARLIER APPLICATION NUMBER: 60/094,657
; EARLIER FILING DATE: 1998-07-30
; NUMBER OF SEQ ID NOS: 1227
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 812
; LENGTH: 97
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (16)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (38)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
US-10-023-282-812

Query Match 42.1%; Score 32; DB 15; Length 97;
Best Local Similarity 39.5%; Pred. No. 3e+02;
Matches 5; Conservative 5; Mismatches 3; Indels 0; Gaps 0;

Qy 3 YDALNVLMMNII 15
Db 66 YDAIAVFLCIHV 78

RESULT 39
US-09-864-761-46653
; Sequence 46653, Application US/09864761
; Patent No. US20020048763A1
; GENERAL INFORMATION:
; APPLICANT: Penn, Sharron G.
; APPLICANT: Hanzel, David K.
; APPLICANT: Chen, Wenheng
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
; FILE REFERENCE: Acomica-X-1
; CURRENT APPLICATION NUMBER: US/09/864,761
; CURRENT FILING DATE: 2001-05-23
; PRIOR APPLICATION NUMBER: US 60/180,312
; PRIOR FILING DATE: 2000-02-04
; PRIOR APPLICATION NUMBER: US 60/207,456
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: US 09/632,366
; PRIOR FILING DATE: 2000-08-03
; PRIOR APPLICATION NUMBER: GB 24263.6
; PRIOR FILING DATE: 2000-10-04
; PRIOR APPLICATION NUMBER: US 60/236,359
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: PCT/US01/00666
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00667
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00664
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00669
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00665
; PRIOR FILING DATE: 2001-01-30

```

```

; PRIOR APPLICATION NUMBER: PCT/US01/00668
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00663
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00662
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00661
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00670
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: US 60/234,687
; PRIOR FILING DATE: 2000-09-21
; PRIOR APPLICATION NUMBER: US 09/608,408
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: US 09/774,203
; PRIOR FILING DATE: 2001-01-29
; NUMBER OF SEQ ID NOS: 49117
; SOFTWARE: Annomax Sequence Listing Engine vers. 1.1
; SEQ ID NO 46653
; LENGTH: 99
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: MAP TO AL031280.6
; OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 1.2
; OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 0.98
; OTHER INFORMATION: SWISSPROT HIT: P25502, EVALUE 6.40e+00
US-09-864-761-46653

Query Match 41.4%; Score 31.5; DB 9; Length 99;
Best Local Similarity 50.0%; Pred. No. 3.7e+02;
Matches 8; Conservative 3; Mismatches 4; Indels 1; Gaps 1;

Qy 1 RYVDALNV-LMAMNII 15
Db 42 RVDVLNVLMQRL 57

RESULT 40
US-09-900-147-7
; Sequence 7, Application US/09900147
; Patent No. US20020103121A1
; GENERAL INFORMATION:
; APPLICANT: La Thangue, Nicholas B
; APPLICANT: Bandara, Lasantha R
; TITLE OF INVENTION: Peptide antagonists of DP transcription factors
; FILE REFERENCE: 620-67
; CURRENT APPLICATION NUMBER: US/09/900,147
; CURRENT FILING DATE: 2001-07-09
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 09/308,935
; PRIOR FILING DATE: EARLIER FILING DATE: 1999-05-27
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: GB 9626589.7
; PRIOR FILING DATE: EARLIER FILING DATE: 1996-12-20
; NUMBER OF SEQ ID NOS: 18
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 7
; LENGTH: 7
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Synthetic
US-09-900-147-7

Query Match 40.8%; Score 31; DB 10; Length 7;
Best Local Similarity 100.0%; Pred. No. 7e+05;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 5 ALNVLMA 11
Db 1 ALNVLMA 7

RESULT 41

```

US-09-966-459A-55  
; Sequence 55, Application US/09966459A  
; Publication No. US2003002237A1  
; GENERAL INFORMATION:  
; APPLICANT: FEDER, J.N.  
; APPLICANT: MINTIER, G.  
; APPLICANT: RAMANATHAN, C.S.  
; APPLICANT: HAWKEN, D.R.  
; APPLICANT: CACACE, A.  
; APPLICANT: BARBER, L.  
; APPLICANT: KORNACKER, M.G.  
; TITLE OF INVENTION: A NOVEL HUMAN G-PROTEIN COUPLED RECEPTOR, HGPRBMV4,  
; TITLE OF INVENTION: EXPRESSED HIGHLY IN PROSTATE, COLON, AND LUNG  
; FILE REFERENCE: D0039NP  
; CURRENT APPLICATION NUMBER: US/09/966.459A  
; CURRENT FILING DATE: 2001-09-26  
; PRIOR APPLICATION NUMBER: 60/235,833  
; PRIOR FILING DATE: 2000-09-27  
; PRIOR APPLICATION NUMBER: 60/261,776  
; PRIOR FILING DATE: 2001-01-16  
; PRIOR APPLICATION NUMBER: 60/305,351  
; PRIOR FILING DATE: 2001-07-13  
; PRIOR APPLICATION NUMBER: 60/313,202  
; PRIOR FILING DATE: 2001-08-17  
; NUMBER OF SEQ ID NOS: 60  
; SOFTWARE: Patent in Ver. 2.1  
; SEQ ID NO 55  
; LENGTH: 15  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Description of Artificial Sequence: Synthetic  
; OTHER INFORMATION: polypeptide  
US-09-966-459A-55

Query Match 40.8%; Score 31; DB 11; Length 15;  
Best Local Similarity 75.0%; Pred. No. 53;  
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Oy 3 YDALNVLN 10  
Db 8 YDALDTLM 15

RESULT 42  
US-09-966-422B-70  
; Sequence 70, Application US/09966422B  
; Publication No. US20030044892A1  
; GENERAL INFORMATION:  
; APPLICANT: Bristol-Myers Squibb Company  
; TITLE OF INVENTION: A NOVEL HUMAN G-PROTEIN COUPLED RECEPTOR, HGPRBMV6, EXPRESSED HIGHLY IN PROSTATE, COLON, AND LUNG  
; TITLE OF INVENTION: SMALL INTESTINE  
; FILE REFERENCE: D0040NP/3053-4119US3  
; CURRENT APPLICATION NUMBER: US/09/966.422B  
; CURRENT FILING DATE: 2002-05-07  
; PRIOR APPLICATION NUMBER: 60/235,602  
; PRIOR FILING DATE: 2000-09-27  
; PRIOR APPLICATION NUMBER: 60/306,604  
; PRIOR FILING DATE: 2001-07-19  
; PRIOR APPLICATION NUMBER: 60/315,412  
; PRIOR FILING DATE: 2001-08-28  
; NUMBER OF SEQ ID NOS: 81  
; SOFTWARE: Patent in version 3.0  
; SEQ ID NO 70  
; LENGTH: 15  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Synthetic Polypeptide  
US-09-966-422B-70

Query Match 40.8%; Score 31; DB 11; Length 15;  
Best Local Similarity 75.0%; Pred. No. 53;

Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;  
Oy 3 YDALNVLN 10  
Db 8 YDALDTLM 15  
RESULT 43  
US-10-262-272A-70  
; Sequence 70, Application US/10262272A  
; Publication No. US20030170671A1  
; GENERAL INFORMATION:  
; APPLICANT: Bristol-Myers Squibb Company  
; TITLE OF INVENTION: NOVEL HUMAN G-PROTEIN COUPLED RECEPTOR, HGPRBMV6  
; FILE REFERENCE: D0044 CIP  
; CURRENT APPLICATION NUMBER: US/10/262.272A  
; CURRENT FILING DATE: 2002-09-27  
; PRIOR APPLICATION NUMBER: U.S. 09/966,422  
; PRIOR FILING DATE: 2001-09-26  
; NUMBER OF SEQ ID NOS: 91  
; SOFTWARE: Patent in version 3.2  
; SEQ ID NO 70  
; LENGTH: 15  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Synthesized Peptide.  
US-10-262-272A-70

Query Match 40.8%; Score 31; DB 12; Length 15;  
Best Local Similarity 75.0%; Pred. No. 53;  
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Oy 3 YDALNVLN 10  
Db 8 YDALDTLM 15

RESULT 44  
US-09-962-756-1214  
; Sequence 1214, Application US/09962756  
; Publication No. US20030195147A1  
; GENERAL INFORMATION:  
; APPLICANT: PILLUTLA, RENUKA  
; APPLICANT: BRISSETTE, RENEE  
; APPLICANT: BLUME, ARTHUR J.  
; APPLICANT: SCHAFER, LAUGE  
; APPLICANT: BRANDT, JAKOB  
; APPLICANT: GOLDSTEIN, NEIL I.  
; APPLICANT: SPETZLER, JANE  
; APPLICANT: OSTERGAARD, SOREN  
; APPLICANT: HANSEN, PER HERTZ  
; TITLE OF INVENTION: INSULIN AND IGF-1 RECEPTOR AGONISTS AND ANTAGONISTS  
; FILE REFERENCE: 1878-4051US1  
; CURRENT APPLICATION NUMBER: US/09/962.756  
; CURRENT FILING DATE: 2001-09-24  
; PRIOR APPLICATION NUMBER: 09/538,038  
; PRIOR FILING DATE: 2000-03-29  
; PRIOR APPLICATION NUMBER: 09/146,127  
; PRIOR FILING DATE: 1998-09-02  
; NUMBER OF SEQ ID NOS: 2227  
; SOFTWARE: Patent in Ver. 2.1  
; SEQ ID NO 1214  
; LENGTH: 19  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Description of Artificial Sequence: Synthetic  
; OTHER INFORMATION: peptide  
US-09-962-756-1214

Query Match 40.8%; Score 31; DB 12; Length 19;  
Best Local Similarity 50.0%; Pred. No. 69;

Tue Feb 17 11:55:52 2004

Matches 6; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

Qy 3 YDALNVLMMNI 14  
| | | | | |  
Db 6 YDAIDRLRMRI 17

RESULT 45  
US-10-253-471-1214  
; Sequence 1214, Application US/10253471  
; Publication No. US20030236190A1  
; GENERAL INFORMATION:  
; APPLICANT: PILIUTLA, RENUKA et al.  
; TITLE OF INVENTION: INSULIN AND IGF-1 RECEPTOR AGONISTS AND ANTAGONISTS  
; FILE REFERENCE: 1878-4057  
; CURRENT APPLICATION NUMBER: US/10/253.471  
; CURRENT FILING DATE: 2002-09-24  
; PRIOR APPLICATION NUMBER: 09/962,756  
; PRIOR FILING DATE: 2001-09-24  
; PRIOR APPLICATION NUMBER: 09/538,038  
; PRIOR FILING DATE: 2000-03-29  
; PRIOR APPLICATION NUMBER: 09/146,127  
; PRIOR FILING DATE: 1998-09-02  
; NUMBER OF SEQ ID NOS: 227  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 1214  
; LENGTH: 19  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Description of Artificial Sequence: Synthetic  
; OTHER INFORMATION: Peptide  
US-10-253-471-1214

Query Match 40.8%; Score 31; DB 12; Length 19;  
Best Local Similarity 50.0%; Pred. No. 69;  
Matches 6; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

Qy 3 YDALNVLMMNI 14  
| | | | | |  
Db 6 YDAIDRLRMRI 17

Search completed: February 17, 2004, 11:02:21  
Job time : 23.4455 secs



GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: February 17, 2004, 10:50:13 ; Search time 11.5644 Seconds  
(without alignments)  
58.540 Million cell updates/sec

Title: US-09-900-147-5

Perfect score: 76

Sequence: 1 RYVDALNVLMMNNIIS 16

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 328717 seqs, 42310858 residues

Total number of hits satisfying chosen parameters: 228043

Minimum DB seq length: 0  
Maximum DB seq length: 100

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Issued Patents AA:\*  
1: /cgn2\_6/prodata/1/iaa/5A COMB.pep.\*  
2: /cgn2\_6/prodata/1/iaa/5B COMB.pep.\*  
3: /cgn2\_6/prodata/1/iaa/6A COMB.pep.\*  
4: /cgn2\_6/prodata/1/iaa/6B COMB.pep.\*  
5: /cgn2\_6/prodata/1/iaa/PCUTS COMB.pep.\*  
6: /cgn2\_6/prodata/1/iaa/backfiles.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	76	100.0	16	3	US-09-308-935-5
2	76	100.0	17	3	US-08-428-131-13
3	76	100.0	17	3	US-09-078-596-13
4	76	100.0	19	3	US-09-308-935-3
5	76	100.0	37	3	US-09-308-935-1
6	76	100.0	72	2	US-08-428-131-11
7	76	100.0	72	3	US-09-078-596-11
8	76	100.0	74	4	US-08-894-139-10
9	68	89.5	19	3	US-09-308-935-15
10	68	89.5	19	3	US-09-308-935-16
11	67	88.2	30	3	US-09-308-935-6
12	64	84.2	14	3	US-09-308-935-11
13	64	84.2	19	3	US-09-308-935-17
14	46	60.5	20	3	US-09-308-935-4
15	44	57.9	11	3	US-09-308-935-9
16	42	55.3	9	3	US-09-308-935-2
17	39	51.3	17	2	US-08-428-131-14
18	39	51.3	17	2	US-09-078-596-14
19	39	51.3	73	2	US-08-428-131-12
20	39	51.3	73	3	US-09-078-596-12
21	39	51.3	74	4	US-08-894-139-5
22	39	51.3	74	4	US-08-894-139-7
23	39	51.3	75	4	US-08-894-139-9
24	38	50.0	69	4	US-08-894-139-8
25	38	50.0	74	4	US-08-894-139-6
26	35	46.1	67	4	US-09-134-001C-4216
27	34	44.7	63	1	US-08-194-338-14

28	34	44.7	88	4	US-08-936-165A-513	Sequence 513, Appl
29	32	42.1	25	1	US-08-378-761A-30	Sequence 30, Appl
30	32	42.1	25	1	US-08-485-286-30	Sequence 30, Appl
31	32	42.1	25	6	5248606-16	Patent No. 5248606
32	32	42.1	97	4	US-09-205-258-812	Sequence 812, Appl
33	31	40.8	7	3	US-09-308-935-7	Sequence 7, Appl
34	31	40.8	22	3	US-08-604-965E-7	Sequence 7, Appl
35	31	40.8	25	3	US-08-604-965E-1	Sequence 1, Appl
36	31	40.8	40	3	US-08-604-965E-2	Sequence 2, Appl
37	31	40.8	42	6	5258287-4	Patent No. 5258287
38	31	40.8	43	4	US-09-217-293-8	Sequence 8, Appl
39	31	40.8	46	3	US-08-740-644-7	Sequence 7, Appl
40	31	40.8	54	4	US-09-205-258-455	Sequence 455, Appl
41	31	40.8	60	6	5258287-1	Patent No. 5258287
42	31	40.8	63	4	US-09-107-532A-5706	Sequence 5706, Ap
43	31	40.8	75	4	US-09-235-451-16	Sequence 16, Appl
44	31	40.8	75	4	US-09-235-451-17	Sequence 17, Appl
45	31	40.8	85	3	US-08-604-965E-9	Sequence 9, Appl

ALIGNMENTS

RESULT 1  
US-09-308-935-5  
; Sequence 5, Application US/09308935  
; Patent No. 6268334  
; GENERAL INFORMATION:  
; APPLICANT: La Thangue, Nicholas B  
; APPLICANT: Bandara, Laseantha R  
; TITLE OF INVENTION: Peptide antagonists of DP transcription factors  
; FILE REFERENCE: 620-67  
; CURRENT APPLICATION NUMBER: US/09/308,935  
; CURRENT FILING DATE: 1999-05-27  
; EARLIER APPLICATION NUMBER: PCT/GB97/03506  
; EARLIER FILING DATE: 1997-12-22  
; EARLIER APPLICATION NUMBER: GB 9626589.7  
; EARLIER FILING DATE: 1996-12-20  
; NUMBER OF SEQ ID NOS: 18  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 5  
; LENGTH: 16  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Description of Artificial Sequence: Synthetic  
US-09-308-935-5

Query Match 100.0%; Score 76; DB 3; Length 16;  
Best Local Similarity 100.0%; Pred. No. 6.7e-08;  
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 RYVDALNVLMMNNIIS 16  
Db 1 RYVDALNVLMMNNIIS 16

RESULT 2  
US-08-428-131-13  
; Sequence 13, Application US/08428131  
; Patent No. 5863757  
; GENERAL INFORMATION:  
; APPLICANT: La Thangue, Nicholas Barrie  
; TITLE OF INVENTION: Transcription Factor DP-1  
; NUMBER OF SEQUENCES: 14  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Nixon & Vanderhye  
; STREET: 1100 No. 5863757th Glebe Road, 8th Floor  
; CITY: Arlington  
; STATE: Virginia  
; COUNTRY: U.S.A.  
; ZIP: 22201-4714  
; COMPUTER READABLE FORM: /

MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25 (EPO)  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/428,131  
FILING DATE: 23-JUN-1995  
CLASSIFICATION: 514  
ATTORNEY/AGENT INFORMATION:  
NAME: Arthur R. Crawford  
REGISTRATION NUMBER: 25,327  
REFERENCE/DOCKET NUMBER: 117-181  
TELEPHONE: (703) 816-4000  
TELEFAX: (703) 816-4100  
INFORMATION FOR SEQ ID NO: 13:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 17 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-428-131-13

Query Match 100.0%; Score 76; DB 2; Length 17;  
Best Local Similarity 100.0%; Pred. No. 7.2e-08;  
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RYVDALNVLMAMNIIS 16  
DB 2 RYVDALNVLMAMNIIS 17

RESULT 3  
US-09-078-596-13  
Sequence 13, Application US/09078596  
Patent No. 6150116  
GENERAL INFORMATION:  
APPLICANT: La Thangue, Nicholas Barrie  
TITLE OF INVENTION: Transcription Factor DP-1  
NUMBER OF SEQUENCES: 14  
CORRESPONDENCE ADDRESSES:  
ADDRESSEE: Nixon & Vanderhye  
STREET: 1100 No. 6150116th Glebe Road, 8th Floor  
CITY: Arlington  
STATE: Virginia  
COUNTRY: U.S.A.  
ZIP: 22201-4714  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25 (EPO)  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/078,596  
FILING DATE:  
CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US/08/428,131  
FILING DATE: 23-JUN-1995  
ATTORNEY/AGENT INFORMATION:  
NAME: Arthur R. Crawford  
REGISTRATION NUMBER: 25,327  
REFERENCE/DOCKET NUMBER: 117-181  
TELEPHONE: (703) 816-4000  
TELEFAX: (703) 816-4100  
INFORMATION FOR SEQ ID NO: 13:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 17 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear

MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25 (EPO)  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/428,131  
FILING DATE: 23-JUN-1995  
CLASSIFICATION: 514  
ATTORNEY/AGENT INFORMATION:  
NAME: Arthur R. Crawford  
REGISTRATION NUMBER: 25,327  
REFERENCE/DOCKET NUMBER: 117-181  
TELEPHONE: (703) 816-4000  
TELEFAX: (703) 816-4100  
INFORMATION FOR SEQ ID NO: 13:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 17 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-09-078-596-13

Query Match 100.0%; Score 76; DB 3; Length 17;  
Best Local Similarity 100.0%; Pred. No. 7.2e-08;  
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RYVDALNVLMAMNIIS 16  
DB 2 RYVDALNVLMAMNIIS 17

RESULT 4  
US-09-308-935-3  
Sequence 3, Application US/09308935  
Patent No. 6268334  
GENERAL INFORMATION:  
APPLICANT: La Thangue, Nicholas B  
APPLICANT: Bandara, Lasaantha R  
TITLE OF INVENTION: Peptide antagonists of DP transcription factors  
FILE REFERENCE: 620-67  
CURRENT APPLICATION NUMBER: US/09/308,935  
CURRENT FILING DATE: 1999-05-27  
PCT/GB97/03506  
EARLIER FILING DATE: 1997-12-22  
EARLIER APPLICATION NUMBER: GB 9626589.7  
EARLIER FILING DATE: 1996-12-20  
NUMBER OF SEQ ID NOS: 18  
SOFTWARE: PatentIn Ver. 2.1  
SEQ ID NO 3  
LENGTH: 19  
TYPE: PRT  
ORGANISM: Artificial Sequence  
FEATURE:  
OTHER INFORMATION: Description of Artificial Sequence: Synthetic  
US-09-308-935-3

Query Match 100.0%; Score 76; DB 3; Length 19;  
Best Local Similarity 100.0%; Pred. No. 8.2e-08;  
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RYVDALNVLMAMNIIS 16  
DB 3 RYVDALNVLMAMNIIS 18

RESULT 5  
US-09-308-935-1  
Sequence 1, Application US/09308935  
Patent No. 6268334  
GENERAL INFORMATION:  
APPLICANT: La Thangue, Nicholas B  
APPLICANT: Bandara, Lasaantha R  
TITLE OF INVENTION: Peptide antagonists of DP transcription factors  
FILE REFERENCE: 620-67  
CURRENT APPLICATION NUMBER: US/09/308,935  
CURRENT FILING DATE: 1999-05-27  
PCT/GB97/03506  
EARLIER FILING DATE: 1997-12-22  
EARLIER APPLICATION NUMBER: GB 9626589.7  
EARLIER FILING DATE: 1996-12-20  
NUMBER OF SEQ ID NOS: 18  
SOFTWARE: PatentIn Ver. 2.1  
SEQ ID NO 1  
LENGTH: 37  
TYPE: PRT  
ORGANISM: Artificial Sequence  
FEATURE:  
OTHER INFORMATION: Description of Artificial Sequence: Synthetic  
US-09-308-935-1

Query Match 100.0%; Score 76; DB 3; Length 37;  
Best Local Similarity 100.0%; Pred. No. 1.8e-07;

Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RYVDALNVLMMNNIIS 16  
|||||  
Db 6 RYVDALNVLMMNNIIS 21

## RESULT 6

US-08-428-131-11  
; Sequence 11, Application US/08428131  
; Patent No. 5863757  
; GENERAL INFORMATION:  
; APPLICANT: La Thangue, Nicholas Barrie  
; TITLE OF INVENTION: Transcription Factor DP-1  
; NUMBER OF SEQUENCES: 14  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Nixon & Vanderhye  
; STREET: 1100 No. 5863757th Glebe Road, 8th Floor  
; CITY: Arlington  
; STATE: Virginia  
; COUNTRY: U.S.A.  
; ZIP: 22201-4714  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent In Release #1.0, Version #1.25 (EPO)  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/428,131  
; FILING DATE: 23-JUN-1995  
; CLASSIFICATION: 514  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Arthur R. Crawford  
; REGISTRATION NUMBER: 25,327  
; REFERENCE/DOCKET NUMBER: 117-181  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (703) 816-4000  
; TELEFAX: (703) 816-4100  
; INFORMATION FOR SEQ ID NO: 11:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 72 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
US-08-428-131-11

Query Match 100.0%; Score 76; DB 2; Length 72;  
Best Local Similarity 100.0%; Pred. No. 4.1e-07;  
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RYVDALNVLMMNNIIS 16  
|||||  
Db 9 RYVDALNVLMMNNIIS 24

## RESULT 7

US-09-078-596-11  
; Sequence 11, Application US/09078596  
; Patent No. 6150116  
; GENERAL INFORMATION:  
; APPLICANT: La Thangue, Nicholas Barrie  
; TITLE OF INVENTION: Transcription Factor DP-1  
; NUMBER OF SEQUENCES: 14  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Nixon & Vanderhye  
; STREET: 1100 No. 6150116th Glebe Road, 8th Floor  
; CITY: Arlington  
; STATE: Virginia  
; COUNTRY: U.S.A.  
; ZIP: 22201-4714  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent In Release #1.0, Version #1.25 (EPO)  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/078,596  
; FILING DATE:  
; CLASSIFICATION:  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US/08/428,131  
; FILING DATE: 23-JUN-1995  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Arthur R. Crawford  
; REGISTRATION NUMBER: 25,327  
; REFERENCE/DOCKET NUMBER: 117-181  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (703) 816-4000  
; TELEFAX: (703) 816-4100  
; INFORMATION FOR SEQ ID NO: 11:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 72 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
US-09-078-596-11

Query Match 100.0%; Score 76; DB 3; Length 72;  
Best Local Similarity 100.0%; Pred. No. 4.1e-07;  
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RYVDALNVLMMNNIIS 16  
|||||  
Db 9 RYVDALNVLMMNNIIS 24

## RESULT 8

US-08-894-139-10  
; Sequence 10, Application US/08894139  
; Patent No. 6448376  
; GENERAL INFORMATION:  
; APPLICANT: LA THANGUE, NICHOLAS B.  
; APPLICANT: BERNARDS, RENE  
; APPLICANT: HUMANS, ELEANORE M.  
; TITLE OF INVENTION: TRANSCRIPTION FACTOR E2F-5  
; NUMBER OF SEQUENCES: 25  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: NIXON & VANDERHYE P.C.  
; STREET: 1100 NORTH GLEBE ROAD  
; CITY: ARLINGTON  
; STATE: VIRGINIA  
; COUNTRY: U.S.A.  
; ZIP: 22201-4714  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent In Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/894,139  
; FILING DATE: 13-AUG-1997  
; CLASSIFICATION: 536  
; ATTORNEY/AGENT INFORMATION:  
; NAME: WILSON, MARY J.  
; REGISTRATION NUMBER: 32,955  
; REFERENCE/DOCKET NUMBER: 620-22  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (703) 816-4000  
; TELEFAX: (703) 816-4100  
; INFORMATION FOR SEQ ID NO: 10:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 74 amino acids  
; TYPE: amino acid  
; STRANDEDNESS:

```
/
/ TOPOLOGY: linear
/ MOLECULE TYPE: peptide
US-08-894-139-10
Query Match 100.0%; Score 76; DB 4; Length 74;
Best Local Similarity 100.0%; Pred. No. 4.3e-07;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RYDNLNVLMMNIIS 16
Db 48 RYDNLNVLMMNIIS 63

RESULT 9
US-09-308-935-15
; Sequence 15, Application US/09308935
; Patent No. 6268334
; GENERAL INFORMATION:
; APPLICANT: La Thangue, Nicholas B
; APPLICANT: Bandara, Lasantha R
; TITLE OF INVENTION: Peptide antagonists of DP transcription factors
; FILE REFERENCE: 620-67
; CURRENT APPLICATION NUMBER: US/09/308,935
; CURRENT FILING DATE: 1999-05-27
; EARLIER APPLICATION NUMBER: PCT/GB97/03506
; EARLIER FILING DATE: 1997-12-22
; EARLIER APPLICATION NUMBER: GB 9626589.7
; EARLIER FILING DATE: 1996-12-20
; NUMBER OF SEQ ID NOS: 18
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 15
; LENGTH: 19
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Mutant peptide
US-09-308-935-15

Query Match 89.5%; Score 68; DB 3; Length 19;
Best Local Similarity 87.5%; Pred. No. 2.5e-06;
Matches 14; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 RYDNLNVLMMNIIS 16
Db 3 RYDNLNVLMMNIIS 18

RESULT 10
US-09-308-935-16
; Sequence 16, Application US/09308935
; Patent No. 6268334
; GENERAL INFORMATION:
; APPLICANT: La Thangue, Nicholas B
; APPLICANT: Bandara, Lasantha R
; TITLE OF INVENTION: Peptide antagonists of DP transcription factors
; FILE REFERENCE: 620-67
; CURRENT APPLICATION NUMBER: US/09/308,935
; CURRENT FILING DATE: 1999-05-27
; EARLIER APPLICATION NUMBER: PCT/GB97/03506
; EARLIER FILING DATE: 1997-12-22
; EARLIER APPLICATION NUMBER: GB 9626589.7
; EARLIER FILING DATE: 1996-12-20
; NUMBER OF SEQ ID NOS: 18
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 16
; LENGTH: 19
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Mutant peptide
US-09-308-935-16

Query Match 89.5%; Score 68; DB 3; Length 19;
Best Local Similarity 87.5%; Pred. No. 2.5e-06;
Matches 14; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 RYDNLNVLMMNIIS 16
Db 3 RYDNLNVLMMNIIS 18

RESULT 11
US-09-308-935-6
; Sequence 6, Application US/09308935
; Patent No. 6268334
; GENERAL INFORMATION:
; APPLICANT: La Thangue, Nicholas B
; APPLICANT: Bandara, Lasantha R
; TITLE OF INVENTION: Peptide antagonists of DP transcription factors
; FILE REFERENCE: 620-67
; CURRENT APPLICATION NUMBER: US/09/308,935
; CURRENT FILING DATE: 1999-05-27
; EARLIER APPLICATION NUMBER: PCT/GB97/03506
; EARLIER FILING DATE: 1997-12-22
; EARLIER APPLICATION NUMBER: GB 9626589.7
; EARLIER FILING DATE: 1996-12-20
; NUMBER OF SEQ ID NOS: 18
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 6
; LENGTH: 30
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Synthetic
US-09-308-935-6

Query Match 88.2%; Score 67; DB 3; Length 30;
Best Local Similarity 100.0%; Pred. No. 6.6e-06;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 YDNLNVLMMNIIS 16
Db 1 YDNLNVLMMNIIS 14

RESULT 12
US-09-308-935-11
; Sequence 11, Application US/09308935
; Patent No. 6268334
; GENERAL INFORMATION:
; APPLICANT: La Thangue, Nicholas B
; APPLICANT: Bandara, Lasantha R
; TITLE OF INVENTION: Peptide antagonists of DP transcription factors
; FILE REFERENCE: 620-67
; CURRENT APPLICATION NUMBER: US/09/308,935
; CURRENT FILING DATE: 1999-05-27
; EARLIER APPLICATION NUMBER: PCT/GB97/03506
; EARLIER FILING DATE: 1997-12-22
; EARLIER APPLICATION NUMBER: GB 9626589.7
; EARLIER FILING DATE: 1996-12-20
; NUMBER OF SEQ ID NOS: 18
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 11
; LENGTH: 14
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Synthetic
US-09-308-935-11

Query Match 84.2%; Score 64; DB 3; Length 14;
Best Local Similarity 100.0%; Pred. No. 9.4e-06;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RYDNLNVLMMNIIS 13
Db 1 RYDNLNVLMMNIIS 13
```

```
Db      2 RYVDALNVLMMN 14

RESULT 13
US-09-308-935-17
; Sequence 17, Application US/09308935
; Patent No. 6268334
; GENERAL INFORMATION:
; APPLICANT: La Thangue, Nicholas B
; APPLICANT: Bandara, Lasantha R
; TITLE OF INVENTION: Peptide antagonists of DP transcription factors
; FILE REFERENCE: 620-67
; CURRENT APPLICATION NUMBER: US/09/308,935
; CURRENT FILING DATE: 1999-05-27
; EARLIER APPLICATION NUMBER: PCT/GB97/03506
; EARLIER FILING DATE: 1997-12-22
; EARLIER APPLICATION NUMBER: GB 9626589.7
; EARLIER FILING DATE: 1996-12-20
; NUMBER OF SEQ ID NOS: 18
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 17
; LENGTH: 19
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Mutant peptide
US-09-308-935-17

Query Match      84.2%; Score 64; DB 3; Length 19;
Best Local Similarity 87.5%; Pred. No. 1.4e-05;
Matches 14; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY      1 RYVDALNVLMMNIIIS 16
        ||||| |||||
Db      3 RYVDARNVMMNIIIS 18

RESULT 14
US-09-308-935-4
; Sequence 4, Application US/09308935
; Patent No. 6268334
; GENERAL INFORMATION:
; APPLICANT: La Thangue, Nicholas B
; APPLICANT: Bandara, Lasantha R
; TITLE OF INVENTION: Peptide antagonists of DP transcription factors
; FILE REFERENCE: 620-67
; CURRENT APPLICATION NUMBER: US/09/308,935
; CURRENT FILING DATE: 1999-05-27
; EARLIER APPLICATION NUMBER: PCT/GB97/03506
; EARLIER FILING DATE: 1997-12-22
; EARLIER APPLICATION NUMBER: GB 9626589.7
; EARLIER FILING DATE: 1996-12-20
; NUMBER OF SEQ ID NOS: 18
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 4
; LENGTH: 20
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Synthetic
US-09-308-935-4

Query Match      60.5%; Score 46; DB 3; Length 20;
Best Local Similarity 100.0%; Pred. No. 0.031;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      7 NVLMAMNIIIS 16
        ||||| |||||
Db      1 NVLMAMNIIIS 10

RESULT 15
US-09-308-935-9
; Sequence 9, Application US/09308935
; Patent No. 6268334
; GENERAL INFORMATION:
; APPLICANT: La Thangue, Nicholas B
; APPLICANT: Bandara, Lasantha R
; TITLE OF INVENTION: Peptide antagonists of DP transcription factors
; FILE REFERENCE: 620-67
; CURRENT APPLICATION NUMBER: US/09/308,935
; CURRENT FILING DATE: 1999-05-27
; EARLIER APPLICATION NUMBER: PCT/GB97/03506
; EARLIER FILING DATE: 1997-12-22
; EARLIER APPLICATION NUMBER: GB 9626589.7
; EARLIER FILING DATE: 1996-12-20
; NUMBER OF SEQ ID NOS: 18
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 9
; LENGTH: 11
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Synthetic
US-09-308-935-9

Query Match      57.9%; Score 44; DB 3; Length 11;
Best Local Similarity 100.0%; Pred. No. 0.036;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 RYVDALNVL 9
        ||||| |||||
Db      3 RYVDALNVL 11

RESULT 16
US-09-308-935-2
; Sequence 2, Application US/09308935
; Patent No. 6268334
; GENERAL INFORMATION:
; APPLICANT: La Thangue, Nicholas B
; APPLICANT: Bandara, Lasantha R
; TITLE OF INVENTION: Peptide antagonists of DP transcription factors
; FILE REFERENCE: 620-67
; CURRENT APPLICATION NUMBER: US/09/308,935
; CURRENT FILING DATE: 1999-05-27
; EARLIER APPLICATION NUMBER: PCT/GB97/03506
; EARLIER FILING DATE: 1997-12-22
; EARLIER APPLICATION NUMBER: GB 9626589.7
; EARLIER FILING DATE: 1996-12-20
; NUMBER OF SEQ ID NOS: 18
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2
; LENGTH: 9
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Synthetic
US-09-308-935-2

Query Match      55.3%; Score 42; DB 3; Length 9;
Best Local Similarity 100.0%; Pred. No. 2.5e+05;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      7 NVLMAMNII 15
        ||||| |||||
Db      1 NVLMAMNII 9

RESULT 17
US-08-428-131-14
; Sequence 14, Application US/08428131
; Patent No. 5863757
; GENERAL INFORMATION:
; APPLICANT: La Thangue, Nicholas Barrie
; TITLE OF INVENTION: Transcription Factor DP-1
```

```

;
; NUMBER OF SEQUENCES: 14
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Nixon & Vanderhye
; STREET: 1100 No. 5863757th Glebe Road, 8th Floor
; CITY: Arlington
; STATE: Virginia
; COUNTRY: U.S.A.
; ZIP: 22201-4714
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25 (EPO)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/428,131
; FILING DATE: 23-JUN-1995
; CLASSIFICATION: 514
; ATTORNEY/AGENT INFORMATION:
; NAME: Arthur R. Crawford
; REGISTRATION NUMBER: 25,327
; REFERENCE/DOCKET NUMBER: 117-181
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703) 816-4100
; TELEFAX: (703) 816-4100
; INFORMATION FOR SEQ ID NO: 14:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 17 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-428-131-14

Query Match 51.3%; Score 39; DB 2; Length 17;
Best Local Similarity 43.8%; Pred. No. 0.51;
Matches 7; Conservative 4; Mismatches 5; Indels 0; Gaps 0;

QY 1 RYVDALNVLMAMNIIS 16
DB 2 RYDITNVLEGIQLIA 17

RESULT 19
US-08-428-131-12
; Sequence 12, Application US/08428131
; Patent No. 5863757
; GENERAL INFORMATION:
; APPLICANT: La Thangue, Nicholas Barrie
; TITLE OF INVENTION: Transcription Factor DP-1
; NUMBER OF SEQUENCES: 14
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Nixon & Vanderhye
; STREET: 1100 No. 5863757th Glebe Road, 8th Floor
; CITY: Arlington
; STATE: Virginia
; COUNTRY: U.S.A.
; ZIP: 22201-4714
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25 (EPO)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/428,131
; FILING DATE: 23-JUN-1995
; CLASSIFICATION: 514
; ATTORNEY/AGENT INFORMATION:
; NAME: Arthur R. Crawford
; REGISTRATION NUMBER: 25,327
; REFERENCE/DOCKET NUMBER: 117-181
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703) 816-4100
; TELEFAX: (703) 816-4100
; INFORMATION FOR SEQ ID NO: 12:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 73 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-428-131-12

Query Match 51.3%; Score 39; DB 2; Length 73;
Best Local Similarity 43.8%; Pred. No. 3;
Matches 7; Conservative 4; Mismatches 5; Indels 0; Gaps 0;

QY 1 RYVDALNVLMAMNIIS 16
DB 9 RYDITNVLEGIQLIA 24

RESULT 20
US-09-078-596-12
; Sequence 12, Application US/09078596
```

/ Patent No. 6150116  
/ GENERAL INFORMATION:  
/ APPLICANT: LA THANGUE, Nicholas Barrie  
/ TITLE OF INVENTION: Transcription Factor DP-1  
/ NUMBER OF SEQUENCES: 14  
/ CORRESPONDENCE ADDRESS:  
/ ADDRESSEE: Nixon & Vanderhye  
/ STREET: 1100 No. 6150116th Glebe Road, 8th Floor  
/ CITY: Arlington  
/ STATE: Virginia  
/ COUNTRY: U.S.A.  
/ ZIP: 22201-4714  
/ COMPUTER READABLE FORM:  
/ MEDIUM TYPE: Floppy disk  
/ COMPUTER: IBM PC compatible  
/ OPERATING SYSTEM: PC-DOS/MS-DOS  
/ SOFTWARE: PatentIn Release #1.0, Version #1.25 (EPO)  
/ CURRENT APPLICATION DATA:  
/ APPLICATION NUMBER: US/09/078,596  
/ FILING DATE:  
/ CLASSIFICATION:  
/ PRIOR APPLICATION DATA:  
/ APPLICATION NUMBER: US/08/428,131  
/ FILING DATE: 23-JUN-1995  
/ ATTORNEY/AGENT INFORMATION:  
/ NAME: Arthur R. Crawford  
/ REGISTRATION NUMBER: 25,327  
/ REFERENCE/DOCKET NUMBER: 117-181  
/ TELECOMMUNICATION INFORMATION:  
/ TELEPHONE: (703) 816-4000  
/ TELEFAX: (703) 816-4100  
/ INFORMATION FOR SEQ ID NO: 12:  
/ SEQUENCE CHARACTERISTICS:  
/ LENGTH: 73 amino acids  
/ TYPE: amino acid  
/ STRANDEDNESS: single  
/ TOPOLOGY: linear  
/ MOLECULE TYPE: protein  
US-09-078-596-12

Query Match 51.3%; Score 39; DB 3; Length 73;  
Best Local Similarity 43.8%; Pred. No. 3;  
Matches 7; Conservative 4; Mismatches 5; Indels 5; Gaps 0;

Qy 1 RYVDALNVLMMNIIS 16  
.:||| : : :  
Db 9 RYDITNVLGIIQLIA 24

RESULT 21  
US-08-894-139-5  
/ Sequence 5, Application US/08894139  
/ Patent No. 6448376  
/ GENERAL INFORMATION:  
/ APPLICANT: LA THANGUE, NICHOLAS B.  
/ APPLICANT: BERNARDS, RENE  
/ APPLICANT: HJLMANS, ELEANORE M.  
/ TITLE OF INVENTION: TRANSCRIPTION FACTOR E2F-5  
/ NUMBER OF SEQUENCES: 25  
/ CORRESPONDENCE ADDRESS:  
/ ADDRESSEE: NIXON & VANDERHYE P.C.  
/ STREET: 1100 NORTH GLEBE ROAD  
/ CITY: ARLINGTON  
/ STATE: VIRGINIA  
/ COUNTRY: U.S.A.  
/ ZIP: 22201-4714  
/ COMPUTER READABLE FORM:  
/ MEDIUM TYPE: Floppy disk  
/ COMPUTER: IBM PC compatible  
/ OPERATING SYSTEM: PC-DOS/MS-DOS  
/ SOFTWARE: PatentIn Release #1.0, Version #1.30  
/ CURRENT APPLICATION DATA:  
/ APPLICATION NUMBER: US/08/894,139

/ FILING DATE: 13-AUG-1997  
/ CLASSIFICATION: 536  
/ ATTORNEY/AGENT INFORMATION:  
/ NAME: WILSON, MARY J.  
/ REGISTRATION NUMBER: 32,955  
/ REFERENCE/DOCKET NUMBER: 620-22  
/ TELECOMMUNICATION INFORMATION:  
/ TELEPHONE: (703) 816-4000  
/ TELEFAX: (703) 816-4100  
/ INFORMATION FOR SEQ ID NO: 5:  
/ SEQUENCE CHARACTERISTICS:  
/ LENGTH: 74 amino acids  
/ TYPE: amino acid  
/ STRANDEDNESS:  
/ TOPOLOGY: linear  
/ MOLECULE TYPE: peptide  
US-08-894-139-5  
Query Match 51.3%; Score 39; DB 4; Length 74;  
Best Local Similarity 43.8%; Pred. No. 3;  
Matches 7; Conservative 4; Mismatches 5; Indels 0; Gaps 0;

Qy 1 RYVDALNVLMMNIIS 16  
.:||| : : :  
Db 47 RYDITNVLGIIQLIA 62

RESULT 22  
US-08-894-139-7  
/ Sequence 7, Application US/08894139  
/ Patent No. 6448376  
/ GENERAL INFORMATION:  
/ APPLICANT: LA THANGUE, NICHOLAS B.  
/ APPLICANT: BERNARDS, RENE  
/ APPLICANT: HJLMANS, ELEANORE M.  
/ TITLE OF INVENTION: TRANSCRIPTION FACTOR E2F-5  
/ NUMBER OF SEQUENCES: 25  
/ CORRESPONDENCE ADDRESS:  
/ ADDRESSEE: NIXON & VANDERHYE P.C.  
/ STREET: 1100 NORTH GLEBE ROAD  
/ CITY: ARLINGTON  
/ STATE: VIRGINIA  
/ COUNTRY: U.S.A.  
/ ZIP: 22201-4714  
/ COMPUTER READABLE FORM:  
/ MEDIUM TYPE: Floppy disk  
/ COMPUTER: IBM PC compatible  
/ OPERATING SYSTEM: PC-DOS/MS-DOS  
/ SOFTWARE: PatentIn Release #1.0, Version #1.30  
/ CURRENT APPLICATION DATA:  
/ APPLICATION NUMBER: US/08/894,139  
/ FILING DATE: 13-AUG-1997  
/ CLASSIFICATION: 536  
/ ATTORNEY/AGENT INFORMATION:  
/ NAME: WILSON, MARY J.  
/ REGISTRATION NUMBER: 32,955  
/ REFERENCE/DOCKET NUMBER: 620-22  
/ TELECOMMUNICATION INFORMATION:  
/ TELEPHONE: (703) 816-4000  
/ TELEFAX: (703) 816-4100  
/ INFORMATION FOR SEQ ID NO: 7:  
/ SEQUENCE CHARACTERISTICS:  
/ LENGTH: 74 amino acids  
/ TYPE: amino acid  
/ STRANDEDNESS:  
/ TOPOLOGY: linear  
/ MOLECULE TYPE: peptide  
US-08-894-139-7

Query Match 51.3%; Score 39; DB 4; Length 74;  
Best Local Similarity 46.7%; Pred. No. 3;  
Matches 7; Conservative 4; Mismatches 4; Indels 0; Gaps 0;

```
QY 1 RYVDALNVLMMNII 15
Db 47 RYDITNVLEGIHLI 61

RESULT 23
US-08-894-139-9
; Sequence 9, Application US/08894139
; Patent No. 6448376
; GENERAL INFORMATION:
; APPLICANT: LA THANGUE, NICHOLAS B.
; APPLICANT: BERNARDS, RENE
; APPLICANT: HIJMAN, ELEANORE M.
; TITLE OF INVENTION: TRANSCRIPTION FACTOR E2F-5
; NUMBER OF SEQUENCES: 25
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: NIXON & VANDERHYE P.C.
; STREET: 1100 NORTH GLEBE ROAD
; CITY: ARLINGTON
; STATE: VIRGINIA
; COUNTRY: U.S.A.
; ZIP: 22201-4714
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/894,139
; FILING DATE: 13-AUG-1997
; CLASSIFICATION: 536
; ATTORNEY/AGENT INFORMATION:
; NAME: WILSON, MARY J.
; REGISTRATION NUMBER: 32,955
; REFERENCE/DOCKET NUMBER: 620-22
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703) 816-4000
; TELEFAX: (703) 816-4100
; INFORMATION FOR SEQ ID NO: 8:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 69 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; US-08-894-139-8
; Query Match 50.0%; Score 38; DB 4; Length 69;
; Best Local Similarity 46.7%; Pred. NO. 4.3;
; Matches 7; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

QY 1 RYVDALNVLMMNII 15
Db 42 RYDITNVLEGIHLI 56

RESULT 25
US-08-894-139-6
; Sequence 6, Application US/08894139
; Patent No. 6448376
; GENERAL INFORMATION:
; APPLICANT: LA THANGUE, NICHOLAS B.
; APPLICANT: BERNARDS, RENE
; APPLICANT: HIJMAN, ELEANORE M.
; TITLE OF INVENTION: TRANSCRIPTION FACTOR E2F-5
; NUMBER OF SEQUENCES: 25
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: NIXON & VANDERHYE P.C.
; STREET: 1100 NORTH GLEBE ROAD
; CITY: ARLINGTON
; STATE: VIRGINIA
; COUNTRY: U.S.A.
; ZIP: 22201-4714
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/894,139
; FILING DATE: 13-AUG-1997
; CLASSIFICATION: 536
; ATTORNEY/AGENT INFORMATION:
; NAME: WILSON, MARY J.
; REGISTRATION NUMBER: 32,955
; REFERENCE/DOCKET NUMBER: 620-22
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703) 816-4000
; TELEFAX: (703) 816-4100
; INFORMATION FOR SEQ ID NO: 6:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 74 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: linear
; US-08-894-139-8
; Query Match 51.3%; Score 39; DB 4; Length 75;
; Best Local Similarity 46.7%; Pred. No. 3.1;
; Matches 7; Conservative 4; Mismatches 4; Indels 0; Gaps 0;

QY 1 RYVDALNVLMMNII 15
Db 48 RYDITNVLEGIHLI 62

RESULT 24
US-08-894-139-8
; Sequence 8, Application US/08894139
; Patent No. 6448376
; GENERAL INFORMATION:
; APPLICANT: LA THANGUE, NICHOLAS B.
; APPLICANT: BERNARDS, RENE
; APPLICANT: HIJMAN, ELEANORE M.
; TITLE OF INVENTION: TRANSCRIPTION FACTOR E2F-5
; NUMBER OF SEQUENCES: 25
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: NIXON & VANDERHYE P.C.
; STREET: 1100 NORTH GLEBE ROAD
; CITY: ARLINGTON
; STATE: VIRGINIA
; COUNTRY: U.S.A.
; ZIP: 22201-4714
```



```
; MOLECULE TYPE: peptide
US-08-894-139-6

Query Match      50.0%; Score 38; DB 4; Length 74;
Best Local Similarity 46.7%; Pred. No. 4.6;
Matches 7; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

QY 1 RYVDALNVLMMNII 15
Db 47 RIYDITNVLGIGLI 61

RESULT 26
US-09-134-001C-4216
; Sequence 4216, Application US/09134001C
; Patent No. 6380370
; GENERAL INFORMATION:
; APPLICANT: Lynn Doucette-Stamm et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO STAPHYLOCOCCUS
; FILE REFERENCE: EPIDERMIDIS FOR DIAGNOSTICS AND THERAPEUTICS
; CURRENT APPLICATION NUMBER: US/09/134,001C
; PRIOR FILING DATE: 1998-08-13
; PRIOR APPLICATION NUMBER: US 60/064,964
; PRIOR FILING DATE: 1997-11-08
; PRIOR APPLICATION NUMBER: US 60/055,779
; PRIOR FILING DATE: 1997-08-14
; NUMBER OF SEQ ID NOS: 5674
; SEQ ID NO 4216
; LENGTH: 67
; TYPE: PRT
; ORGANISM: Staphylococcus epidermidis
US-09-134-001C-4216

Query Match      46.1%; Score 35; DB 4; Length 67;
Best Local Similarity 40.8%; Pred. No. 15;
Matches 6; Conservative 6; Mismatches 3; Indels 0; Gaps 0;

QY 2 VYDALNVLMMNIIIS 16
Db 11 LYSKLNILMSIILN 25

RESULT 27
US-08-194-338-14
; Sequence 14, Application US/08194338
; Patent No. 5474898
; GENERAL INFORMATION:
; APPLICANT: Venter, John C.
; APPLICANT: Fraser, Claire M.
; APPLICANT: McCombie, William R.
; TITLE OF INVENTION: OCTOPAMINE RECEPTOR
; NUMBER OF SEQUENCES: 16
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Knobbe, Martens, Olson and Bear
; STREET: 620 Newport Center Drive, Sixteenth Floor
; CITY: Newport Beach
; STATE: CA
; COUNTRY: USA
; ZIP: 92660
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/194,338
; FILING DATE: 08-FEB-1994
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; PRIOR APPLICATION NUMBER: US 07/676,174
; FILING DATE: 28-MAR-1991
; ATTORNEY/AGENT INFORMATION:
```

```
; NAME: Israelsen, Ned A.
; REGISTRATION NUMBER: 29,655
; REFERENCE/DOCKET NUMBER: NIH101.001DV1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (619) 235-8550
; TELEFAX: (619) 235-0176
; INFORMATION FOR SEQ ID NO: 14:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 63 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; FRAGMENT TYPE: internal
US-08-194-338-14

Query Match      44.7%; Score 34; DB 1; Length 63;
Best Local Similarity 60.0%; Pred. No. 21;
Matches 6; Conservative 4; Mismatches 0; Indels 0; Gaps 0;

QY 7 NVLMNIIIS 16
Db 3 NVLVALNILA 12

RESULT 28
US-08-936-165A-513
; Sequence 513, Application US/08936165A
; Patent No. 6348582
; GENERAL INFORMATION:
; APPLICANT: Black, Michael
; APPLICANT: Burnham, Martin
; APPLICANT: Hodgson, John
; APPLICANT: Knowles, David
; APPLICANT: Lonetto, Michael
; APPLICANT: Nicholas, Richard
; APPLICANT: Pratt, Julie
; APPLICANT: Reichard, Richard
; APPLICANT: Rosenberg, Martin
; APPLICANT: Ward, Judith
; TITLE OF INVENTION: No. 6348582el Prokaryotic Polynucleotides,
; NUMBER OF SEQUENCES: 534
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: SmithKline Beecham Corporation
; STREET: 709 Swedeland Road
; CITY: King of Prussia
; STATE: PA
; COUNTRY: USA
; ZIP: 19406-0939
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/936,165A
; FILING DATE: 24-SEP-1997
; CLASSIFICATION: 536
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 60/027,032
; FILING DATE: 24-SEP-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Gimmi, Edward R.
; REGISTRATION NUMBER: 38,891
; REFERENCE/DOCKET NUMBER: P50549
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 610-270-4478
; TELEFAX: 610-270-5090
; TELEX:
; INFORMATION FOR SEQ ID NO: 513:
```

```
; SEQUENCE CHARACTERISTICS:
; LENGTH: 88 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: Protein
US-08-936-165A-513

Query Match 44.7%; Score 34; DB 4; Length 88;
Best Local Similarity 56.2%; Pred. No. 32;
Matches 9; Conservative 3; Mismatches 2; Indels 2; Gaps 1;

QY 2 VYDAL--NVLMMNVII 15
Db 48 MYDAKMDNVLVPINII 63

RESULT 29
US-08-378-761A-30
; Sequence 30, Application US/08378761A
; Patent No. 5635384
; GENERAL INFORMATION:
; APPLICANT: WALSH, TERENCE A
; APPLICANT: HEY, TIMOTHY D
; TITLE OF INVENTION: RIBOSOME-INACTIVATING PROTEINS, INACTIVE
; TITLE OF INVENTION: PRECURSOR FORMS THEREOF, A PROCESS FOR MAKING A METHOD OF
; TITLE OF INVENTION: USING
; NUMBER OF SEQUENCES: 81
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: ANDREA T. BORUCKI
; STREET: 9330 ZIONSVILLE ROAD
; CITY: INDIANAPOLIS
; STATE: IN
; COUNTRY: US
; ZIP: 46268
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent in Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/485,286
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/378761
; FILING DATE: 26-JAN-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: BORUCKI, ANDREA T
; REGISTRATION NUMBER: 33651
; REFERENCE/DOCKET NUMBER: 38272B
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (317) 337-4846
; INFORMATION FOR SEQ ID NO: 30:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 25 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-485-286-30

Query Match 42.1%; Score 32; DB 1; Length 25;
Best Local Similarity 55.6%; Pred. No. 16;
Matches 5; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 1 RYVDALNVL 9
Db 11 KTYDSLNI 19

RESULT 31
5248606-16
; Patent No. 5248606
; APPLICANT: WALSH, TERENCE A.; HEY, TIMOTHY D.; MORGAN,
; ALICE E.R.
; TITLE OF INVENTION: DNA ENCODING INACTIVE PRECURSOR AND
; ACTIVE FORMS OF MAIZE RIBOSOME INACTIVATIN
; NUMBER OF SEQUENCES: 49
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/535,636
; FILING DATE: 11-JUN-1990
; SEQ ID NO:16:
; LENGTH: 25
5248606-16

Query Match 42.1%; Score 32; DB 6; Length 25;
Best Local Similarity 55.6%; Pred. No. 16;
Matches 5; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 1 RYVDALNVL 9
Db 11 KTYDSLNI 19

RESULT 30
US-08-485-286-30
; Sequence 30, Application US/08485286
; Patent No. 5646026
```

Db 11 KTYDSLNI 19

RESULT 32

US-09-205-258-812

Sequence 812, Application US/09205258

Patent No. 625174

GENERAL INFORMATION:

APPLICANT: Young et al.

TITLE OF INVENTION: 207 Human Secreted Proteins

FILE REFERENCE: P2007P1

CURRENT APPLICATION NUMBER: US/09/205,258

CURRENT FILING DATE: 1998-12-04

EARLIER APPLICATION NUMBER: PCT/US98/11422

EARLIER FILING DATE: 1998-06-04

EARLIER APPLICATION NUMBER: 60/048,885

EARLIER FILING DATE: 1997-06-06

EARLIER APPLICATION NUMBER: 60/049,375

EARLIER FILING DATE: 1997-06-06

EARLIER APPLICATION NUMBER: 60/048,881

EARLIER FILING DATE: 1997-06-06

EARLIER APPLICATION NUMBER: 60/048,880

EARLIER FILING DATE: 1997-06-06

EARLIER APPLICATION NUMBER: 60/048,896

EARLIER FILING DATE: 1997-06-06

EARLIER APPLICATION NUMBER: 60/049,020

EARLIER FILING DATE: 1997-06-06

EARLIER APPLICATION NUMBER: 60/048,876

EARLIER FILING DATE: 1997-06-06

EARLIER APPLICATION NUMBER: 60/048,895

EARLIER FILING DATE: 1997-06-06

EARLIER APPLICATION NUMBER: 60/048,884

EARLIER FILING DATE: 1997-06-06

EARLIER APPLICATION NUMBER: 60/048,894

EARLIER FILING DATE: 1997-06-06

EARLIER APPLICATION NUMBER: 60/048,971

EARLIER FILING DATE: 1997-06-06

EARLIER APPLICATION NUMBER: 60/048,964

EARLIER FILING DATE: 1997-06-06

EARLIER APPLICATION NUMBER: 60/048,882

EARLIER FILING DATE: 1997-06-06

EARLIER APPLICATION NUMBER: 60/048,899

EARLIER FILING DATE: 1997-06-06

EARLIER APPLICATION NUMBER: 60/048,893

EARLIER FILING DATE: 1997-06-06

EARLIER APPLICATION NUMBER: 60/048,900

EARLIER FILING DATE: 1997-06-06

EARLIER APPLICATION NUMBER: 60/048,901

EARLIER FILING DATE: 1997-06-06

EARLIER APPLICATION NUMBER: 60/048,892

EARLIER FILING DATE: 1997-06-06

EARLIER APPLICATION NUMBER: 60/048,915

EARLIER FILING DATE: 1997-06-06

EARLIER APPLICATION NUMBER: 60/049,019

EARLIER FILING DATE: 1997-06-06

EARLIER APPLICATION NUMBER: 60/048,970

EARLIER FILING DATE: 1997-06-06

EARLIER APPLICATION NUMBER: 60/048,972

EARLIER FILING DATE: 1997-06-06

EARLIER APPLICATION NUMBER: 60/048,916

EARLIER FILING DATE: 1997-06-06

EARLIER APPLICATION NUMBER: 60/049,373

EARLIER FILING DATE: 1997-06-06

EARLIER APPLICATION NUMBER: 60/048,875

EARLIER FILING DATE: 1997-06-06

EARLIER APPLICATION NUMBER: 60/049,374

EARLIER FILING DATE: 1997-06-06

EARLIER APPLICATION NUMBER: 60/048,917

EARLIER FILING DATE: 1997-06-06

EARLIER APPLICATION NUMBER: 60/048,949

EARLIER FILING DATE: 1997-06-06

EARLIER APPLICATION NUMBER: 60/048,974

EARLIER FILING DATE: 1997-06-06

EARLIER APPLICATION NUMBER: 60/048,883

EARLIER FILING DATE: 1997-06-06

EARLIER APPLICATION NUMBER: 60/048,897

EARLIER FILING DATE: 1997-06-06

EARLIER APPLICATION NUMBER: 60/048,898

EARLIER FILING DATE: 1997-06-06

EARLIER APPLICATION NUMBER: 60/048,962

EARLIER FILING DATE: 1997-06-06

EARLIER APPLICATION NUMBER: 60/048,963

EARLIER FILING DATE: 1997-06-06

EARLIER APPLICATION NUMBER: 60/048,877

EARLIER FILING DATE: 1997-06-06

EARLIER APPLICATION NUMBER: 60/048,878

EARLIER FILING DATE: 1997-06-06

EARLIER APPLICATION NUMBER: 60/070,923

EARLIER FILING DATE: 1997-12-18

EARLIER APPLICATION NUMBER: 60/092,921

EARLIER FILING DATE: 1998-07-15

EARLIER APPLICATION NUMBER: 60/094,657

EARLIER FILING DATE: 1998-07-30

NUMBER OF SEQ ID NOS: 1227

SOFTWARE: PatentIn Ver. 2.0

SEQ ID NO 812

LENGTH: 97

TYPE: PRT

ORGANISM: Homo sapiens

FEATURE:

NAME/KEY: SITE

LOCATION: (16)

OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids

FEATURE:

NAME/KEY: SITE

LOCATION: (38)

OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids

US-09-205-258-812

Query Match 42.1%; Score 32; DB 4; Length 97;

Best Local Similarity 38.5%; Pred. No. 83;

Matches 5; Conservative 5; Mismatches 3; Indels 0; Gaps 0;

QY 3 YDALNVLMMNII 15

DB 66 YDAIVFLCIHIV 78

RESULT 33

US-09-308-935-7

Sequence 7, Application US/09308935

Patent No. 6268334

GENERAL INFORMATION:

APPLICANT: La Thangue, Nicholas B

APPLICANT: Bandara, Lasantha R

TITLE OF INVENTION: Peptide antagonists of DP transcription factors

FILE REFERENCE: 620-67

CURRENT APPLICATION NUMBER: US/09/308,935

CURRENT FILING DATE: 1999-05-27

EARLIER APPLICATION NUMBER: PCT/GB97/03506

EARLIER FILING DATE: 1997-12-22

EARLIER APPLICATION NUMBER: GB 9626589.7

EARLIER FILING DATE: 1996-12-20

NUMBER OF SEQ ID NOS: 18

SOFTWARE: PatentIn Ver. 2.1

SEQ ID NO 7

LENGTH: 7

TYPE: PRT

ORGANISM: Artificial Sequence

FEATURE:

OTHER INFORMATION: Description of Artificial Sequence: Synthetic

US-09-308-935-7

Query Match 40.8%; Score 31; DB 3; Length 7;

Best Local Similarity 100.0%; Pred. No. 2.5e+05;

Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

us-09-900-147-5.ra1

Tue Feb 17 11:55:52 2004

QY 5 ALNVLMA 11  
DB 1 ALNVLMA 7

RESULT 34  
US-08-604-965E-7  
; Sequence 7, Application US/08604965E  
; Patent No. 6046033  
; GENERAL INFORMATION:  
; APPLICANT: Goto, Masaaki  
; APPLICANT: Tsuda, Eisuke  
; APPLICANT: Yano, Kazuki  
; APPLICANT: Kobayashi, Fumie  
; APPLICANT: Yamaguchi, Kyoji  
; APPLICANT: Washida, Naohiro  
; APPLICANT: Satake, Toshiko  
; APPLICANT: Morinaga, Tomonori  
; APPLICANT: Ueda, Masatsugu  
; APPLICANT: Higashio, Kanji  
; TITLE OF INVENTION: Basic Osteoblast Growth FactorII (BOGF-  
; TITLE OF INVENTION: II)  
; NUMBER OF SEQUENCES: 20  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Burgess, Ryan and Wayne  
; STREET: 370 Lexington Avenue, Suite 2105  
; CITY: NEW YORK  
; STATE: NEW YORK  
; COUNTRY: UNITED STATES OF AMERICA  
; ZIP: 10017  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: 5 1/4 inch diskette  
; COMPUTER: PC'S LIMITED SYSTEM 200  
; OPERATING SYSTEM: DOS  
; SOFTWARE: WORDPERFECT 5.1  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/604,965E  
; FILING DATE:  
; CLASSIFICATION: 435  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: PCT/JP95/01270  
; FILING DATE: June 27, 1994  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Wayne, Milton J.  
; REGISTRATION NUMBER: 17,906  
; REFERENCE/DOCKET NUMBER: U-Wp-5212  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 212-683-8150  
; TELEFAX: 212-532-4285  
; TELEX: 423794  
; INFORMATION FOR SEQ ID NO: 1:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 25  
; TYPE: amino acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: peptide  
; US-08-604-965E-1

Query Match 40.8%; Score 31; DB 3; Length 25;  
Best Local Similarity 46.2%; Pred. No. 25;  
Matches 6; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

QY 4 DALNVLMMNNIIS 16  
DB 13 DTLNHLKFLNVL 25

RESULT 36  
US-08-604-965E-2  
; Sequence 2, Application US/08604965E  
; Patent No. 6046033  
; GENERAL INFORMATION:  
; APPLICANT: Goto, Masaaki  
; APPLICANT: Tsuda, Eisuke  
; APPLICANT: Yano, Kazuki  
; APPLICANT: Kobayashi, Fumie  
; APPLICANT: Yamaguchi, Kyoji  
; APPLICANT: Washida, Naohiro  
; APPLICANT: Satake, Toshiko  
; APPLICANT: Morinaga, Tomonori  
; APPLICANT: Ueda, Masatsugu  
; APPLICANT: Higashio, Kanji

QY 4 DALNVLMMNNIIS 16  
DB 1 DTLNHLKFLNVL 13

RESULT 35  
US-08-604-965E-1  
; Sequence 1, Application US/08604965E  
; Patent No. 6046033  
; GENERAL INFORMATION:  
; APPLICANT: Goto, Masaaki  
; APPLICANT: Tsuda, Eisuke  
; APPLICANT: Yano, Kazuki  
; APPLICANT: Kobayashi, Fumie  
; APPLICANT: Yamaguchi, Kyoji  
; APPLICANT: Washida, Naohiro  
; APPLICANT: Satake, Toshiko  
; APPLICANT: Morinaga, Tomonori  
; APPLICANT: Ueda, Masatsugu  
; APPLICANT: Higashio, Kanji

```
; TITLE OF INVENTION: Basic Osteoblast Growth FactorII (bOGF-
; TITLE OF INVENTION: II)
; NUMBER OF SEQUENCES: 20
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Burgess, Ryan and Wayne
; STREET: 370 Lexington Avenue, Suite 2105
; CITY: NEW YORK
; STATE: NEW YORK
; COUNTRY: UNITED STATES OF AMERICA
; ZIP: 10017
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 5 1/4 inch diskette
; COMPUTER: PC'S LIMITED SYSTEM 200
; OPERATING SYSTEM: DOS
; SOFTWARE: WORDPERFECT 5.1
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/604,965E
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/JP95/01270
; FILING DATE: June 27, 1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Wayne, Milton J.
; REGISTRATION NUMBER: 17,906
; REFERENCE/DOCKET NUMBER: U-WP-5212
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 212-683-8150
; TELEFAX: 212-532-4285
; TELEX: 423794
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 40
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; US-08-604-965E-2

Query Match 40.8%; Score 31; DB 3; Length 40;
Best Local Similarity 46.2%; Pred. No. 43;
Matches 6; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

QY 4 DALNLVLMAMNIIIS 16
Db 13 DTLNHLKFLNVL 25

RESULT 37
5258287-4
; Patent No. 5258287
; APPLICANT: BAXTER, ROBERT C.; WOOD, WILLIAM I.
; TITLE OF INVENTION: DNA ENCODING AND METHODS OF PRODUCTION
; OF INSULIN-LIKE GROWTH FACTOR BINDING PROTEIN BP53
; NUMBER OF SEQUENCES: 58
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/171,623
; FILING DATE: 22-MAR-1988
; SEQ ID NO: 4:
; LENGTH: 42
5258287-4

Query Match 40.8%; Score 31; DB 6; Length 42;
Best Local Similarity 46.2%; Pred. No. 46;
Matches 6; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

QY 4 DALNLVLMAMNIIIS 16
Db 14 DTLNHLKFLNVL 26

RESULT 38
US-09-217-293-8
```

```
; Sequence 8, Application US/09217293
; Patent No. 6337181
; GENERAL INFORMATION:
; APPLICANT: Stewart, Jeffrey J
; TITLE OF INVENTION: A METHOD OF SPECIFYING VACCINE COMPONENTS FOR VIRAL
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: QUNASISPECIES
; STREET:
; CITY:
; STATE:
; COUNTRY:
; ZIP:
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.5 inch, 1.4 Mb storage
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: MS-DOS/ Microsoft Windows 3.1
; SOFTWARE: Wordperfect for Windows 5.1
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/740,644
; FILING DATE: October 31, 1996
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 60/006,168
; FILING DATE: No. 603877ember 2, 1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Nelson, M. Bud
; REGISTRATION/DOCKET NUMBER: 11520.0065
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (716) 856-4000
; TELEFAX: (716) 849-0349
; INFORMATION FOR SEQ ID NO: 7:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 46 amino acid residues
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE:
; DESCRIPTION: protein

US-09-217-293-8

Query Match 40.8%; Score 31; DB 4; Length 43;
Best Local Similarity 63.6%; Pred. No. 47;
Matches 7; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 3 YDALNLVLMAMN 13
Db 8 YPALNVTMPNN 18

RESULT 39
US-08-740-644-7
; Sequence 7, Application US/08740644
; Patent No. 6033877
; GENERAL INFORMATION:
; APPLICANT: Timothy F. Murphy
; APPLICANT: Kyungcheol Yi
; TITLE OF INVENTION: Peptide Expression And Delivery System
; NUMBER OF SEQUENCES: 8
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Hodgson, Russ, Andrews, Woods & Goodyear
; STREET: 1800 One M&T Plaza
; CITY: Buffalo
; STATE: New York
; COUNTRY: United States
; ZIP: 14203-2391
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.5 inch, 1.4 Mb storage
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: MS-DOS/ Microsoft Windows 3.1
; SOFTWARE: Wordperfect for Windows 5.1
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/740,644
; FILING DATE: October 31, 1996
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 60/006,168
; FILING DATE: No. 603877ember 2, 1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Nelson, M. Bud
; REGISTRATION/DOCKET NUMBER: 11520.0065
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (716) 856-4000
; TELEFAX: (716) 849-0349
; INFORMATION FOR SEQ ID NO: 7:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 46 amino acid residues
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE:
; DESCRIPTION: protein
```

```

; ORIGINAL SOURCE:
; ORGANISM: H. influenzae
US-08-740-644-7

Query Match      40.8%; Score 31; DB 3; Length 46;
Best Local Similarity 50.0%; Pred. No. 51;
Matches 5; Conservative 4; Mismatches 1; Indels 0; Gaps 0;

Qy      3 YDALNVLMM 12
Db      7 YDANNIIVAI 16
||| |::|:|
||| |::|:|

RESULT 40
US-09-205-258-455
; Sequence 455, Application US/09205258
; Patent No. 6525174
; GENERAL INFORMATION:
; APPLICANT: Young et al.
; TITLE OF INVENTION: 207 Human Secreted Proteins
; FILE REFERENCE: P2007P1
; CURRENT APPLICATION NUMBER: US/09/205,258
; CURRENT FILING DATE: 1998-12-04
; EARLIER APPLICATION NUMBER: PCT/US98/11422
; EARLIER FILING DATE: 1998-06-04
; EARLIER APPLICATION NUMBER: 60/048,885
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/049,375
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,881
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,880
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,896
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/049,020
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,876
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,895
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,884
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,894
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,971
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,964
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,882
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,899
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,893
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,900
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,901
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,892
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,915
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/049,019
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,970
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,972
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,916
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/049,373
; EARLIER FILING DATE: 1997-06-06

; EARLIER APPLICATION NUMBER: 60/048,875
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/049,374
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,917
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,949
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,974
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,883
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,897
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,898
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,962
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,963
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,877
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,878
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/070,923
; EARLIER FILING DATE: 1997-12-18
; EARLIER APPLICATION NUMBER: 60/092,921
; EARLIER FILING DATE: 1998-07-15
; EARLIER APPLICATION NUMBER: 60/094,657
; EARLIER FILING DATE: 1998-07-30
; NUMBER OF SEQ ID NOS: 1227
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 455
; LENGTH: 54
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-205-258-455

Query Match      40.8%; Score 31; DB 4; Length 54;
Best Local Similarity 37.5%; Pred. No. 63;
Matches 6; Conservative 5; Mismatches 5; Indels 0; Gaps 0;

Qy      1 RYVDALNVLMMNIIS 16
Db      35 RYVPAMHFTLCVHIYS 50
||| |::|:|
||| |::|:|

RESULT 41
5258287-1
; Patent No. 5258287
; APPLICANT: BAXTER, ROBERT C.; WOOD, WILLIAM I.
; TITLE OF INVENTION: DNA ENCODING AND METHODS OF PRODUCTION
; OF INSULIN-LIKE GROWTH FACTOR BINDING PROTEIN BP53
; NUMBER OF SEQUENCES: 58
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/171,623
; FILING DATE: 22-MAR-1988
; SEQ ID NO:1:
; LENGTH: 60
5258287-1

Query Match      40.8%; Score 31; DB 6; Length 60;
Best Local Similarity 46.2%; Pred. No. 71;
Matches 6; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

Qy      4 DALNVLMMNIIS 16
Db      14 DTLNHLKFLNVLIS 26
||| |::|:|
||| |::|:|

RESULT 42
US-09-107-532A-5706
; Sequence 5706, Application US/09107532A
```

Patent No. 6583275  
GENERAL INFORMATION:  
APPLICANT: Lynn A Doucette-Stamm and David Bush  
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO  
NUMBER OF SEQUENCES: 7310  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: GENOME THERAPEUTICS CORPORATION  
STREET: 100 Beaver Street  
CITY: Waltham  
STATE: Massachusetts  
COUNTRY: USA  
ZIP: 02354  
COMPUTER READABLE FORM:  
MEDIUM TYPE: CD-ROM ISO9660  
COMPUTER: PC  
OPERATING SYSTEM: <Unknown>  
SOFTWARE: ASCII  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/107,532A  
FILING DATE: 30-Jun-1998  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 60/085,598  
FILING DATE: 14 May 1998  
APPLICATION NUMBER: 60/051571  
FILING DATE: July 2, 1997  
ATTORNEY/AGENT INFORMATION:  
NAME: Ariniello, Pamela Deneke  
REGISTRATION NUMBER: 40,489  
REFERENCE/DOCKET NUMBER: GTC-012  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (781)893-5007  
TELEFAX: (781)893-8277  
INFORMATION FOR SEQ ID NO: 5706:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 63 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
HYPOTHETICAL: YES  
ORIGINAL SOURCE:  
ORGANISM: Enterococcus faecium  
FEATURE:  
NAME/KEY: misc feature  
LOCATION: (B) LOCATION 1..63  
SEQUENCE DESCRIPTION: SEQ ID NO: 5706:  
US-09-107-532A-5706

Query Match 40.8%; Score 31; DB 4; Length 63;  
Best Local Similarity 66.7%; Pred. No. 75;  
Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 RYVDALNVL 9  
| | | | | : |  
Db 40 RYVDFLEIL 48

RESULT 43  
US-09-235-451-16  
Sequence 16, Application US/09235451  
GENERAL INFORMATION:  
APPLICANT: Julius, David J.  
APPLICANT: Caterina, Michael J.  
APPLICANT: Brake, Anthony J.  
TITLE OF INVENTION: NUCLEIC ACID SEQUENCES ENCODING  
TITLE OF INVENTION: CAPSAICIN RECEPTOR AND CAPSAICIN RECEPTOR-RELATED  
TITLE OF INVENTION: POLYPEPTIDES AND USES THEREOF  
FILE REFERENCE: 9076/084CIP  
CURRENT APPLICATION NUMBER: US/09/235,451  
CURRENT FILING DATE: 1999-01-22  
PRIOR APPLICATION NUMBER: 60/072,151  
PRIOR FILING DATE: 1998-01-22  
PRIOR APPLICATION NUMBER: 08/915,461

PRIOR FILING DATE: 1997-08-20  
NUMBER OF SEQ ID NOS: 48  
SOFTWARE: FastSEQ for Windows Version 3.0  
SEQ ID NO: 16  
LENGTH: 75  
TYPE: PRT  
ORGANISM: Caliphora  
US-09-235-451-16  
Query Match 40.8%; Score 31; DB 4; Length 75;  
Best Local Similarity 23.1%; Pred. No. 93;  
Matches 3; Conservative 7; Mismatches 3; Indels 0; Gaps 0;  
QY 3 YDALNVLMMNII 15  
| | | | | : |  
Db 35 YSVINIIVLLNML 47  
RESULT 44  
US-09-235-451-17  
Sequence 17, Application US/09235451  
GENERAL INFORMATION:  
APPLICANT: Julius, David J.  
APPLICANT: Caterina, Michael J.  
TITLE OF INVENTION: NUCLEIC ACID SEQUENCES ENCODING  
TITLE OF INVENTION: CAPSAICIN RECEPTOR AND CAPSAICIN RECEPTOR-RELATED  
TITLE OF INVENTION: POLYPEPTIDES AND USES THEREOF  
FILE REFERENCE: 9076/084CIP  
CURRENT APPLICATION NUMBER: US/09/235,451  
CURRENT FILING DATE: 1999-01-22  
PRIOR APPLICATION NUMBER: 60/072,151  
PRIOR FILING DATE: 1998-01-22  
PRIOR APPLICATION NUMBER: 08/915,461  
PRIOR FILING DATE: 1997-08-20  
NUMBER OF SEQ ID NOS: 48  
SOFTWARE: FastSEQ for Windows Version 3.0  
SEQ ID NO: 17  
LENGTH: 75  
TYPE: PRT  
ORGANISM: D. melanogaster  
US-09-235-451-17

Query Match 40.8%; Score 31; DB 4; Length 75;  
Best Local Similarity 23.1%; Pred. No. 93;  
Matches 3; Conservative 7; Mismatches 3; Indels 0; Gaps 0;

QY 3 YDALNVLMMNII 15  
| | | | | : |  
Db 35 YSVINIIVLLNML 47

RESULT 45  
US-08-604-965E-9  
Sequence 9, Application US/08604965E  
Patent No. 6046033  
GENERAL INFORMATION:  
APPLICANT: Goto, Masaaki  
APPLICANT: Tsuda, Eisuke  
APPLICANT: Yano, Kazuki  
APPLICANT: Kobayashi, Fumie  
APPLICANT: Yamaguchi, Kyoji  
APPLICANT: Washida, Naohiro  
APPLICANT: Satake, Toshiko  
APPLICANT: Morinaga, Tomonori  
APPLICANT: Ueda, Masatsugu  
APPLICANT: Higashio, Kanji  
TITLE OF INVENTION: Basic Osteoblast Growth FactorII (bOGF-  
TITLE OF INVENTION: II)  
NUMBER OF SEQUENCES: 20  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Burgess, Ryan and Wayne  
STREET: 370 Lexington Avenue, Suite 2105

Tue Feb 17 11:55:52 2004

CITY: NEW YORK  
STATE: NEW YORK  
COUNTRY: UNITED STATES OF AMERICA  
ZIP: 10017  
COMPUTER READABLE FORM:  
MEDIUM TYPE: 5 1/4 inch diskette  
COMPUTER: PC'S LIMITED SYSTEM 200  
OPERATING SYSTEM: DOS  
SOFTWARE: WORDPERFECT 5.1  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/604,965E  
FILING DATE:  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: PCT/JP95/01270  
FILING DATE: June 27, 1994  
ATTORNEY/AGENT INFORMATION:  
NAME: Wayne, Milton J.  
REGISTRATION NUMBER: 17,906  
REFERENCE/DOCKET NUMBER: U-Wp-5212  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 212-683-8150  
TELEFAX: 212-532-4285  
TELEX: 423794  
INFORMATION FOR SEQ ID NO: 9:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 85  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-604-965E-9

Query Match 40.8%; Score 31; DB 3; Length 85;  
Best Local Similarity 46.2%; Pred. No. 1.1e+02;  
Matches 6; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

Qy 4 DALNVLMMNNIIS 16  
Db 13 DTLNHLKFLNVLIS 25

Search completed: February 17, 2004, 10:59:41  
Job time : 12.5644 secs



GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: February 17, 2004, 10:50:12 ; Search time 22.6535 Seconds  
(without alignments)  
182.261 Million cell updates/sec

Title: US-09-900-147-5

Perfect score: 76

Sequence: 1 RVDALNVLMMNIIS 16

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 830525 seqs, 258052604 residues

Total number of hits satisfying chosen parameters: 146963

Minimum DB seq length: 0

Maximum DB seq length: 100

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

SPTREMBL\_23.\*

- 1: sp\_archaea.\*
- 2: sp\_bacteria.\*
- 3: sp\_fungi.\*
- 4: sp\_human.\*
- 5: sp\_invertebrate.\*
- 6: sp\_mammal.\*
- 7: sp\_mhc.\*
- 8: sp\_organelle.\*
- 9: sp\_phase.\*
- 10: sp\_plant.\*
- 11: sp\_rodent.\*
- 12: sp\_virus.\*
- 13: sp\_vertebrate.\*
- 14: sp\_unclassified.\*
- 15: sp\_rvirus.\*
- 16: sp\_bacteriap.\*
- 17: sp\_archaeap.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match %	Length	ID	Description
1	35	46.1	68	16 Q8DSX0	Q8dsx0 streptococ
2	35	46.1	80	11 Q8C8H4	Q8c8h4 mus musculu
3	34	44.7	74	16 Q9VZ55	Q9vz55 staphylococ
4	34	44.7	80	11 Q99MV9	Q99mv9 mus musculu
5	33	43.4	66	17 Q9HQ82	Q9hq82 halobacteri
6	33	43.4	82	9 Q48472	Q48472 bacterioph
7	32	42.1	55	8 P92499	P92499 apis mellif
8	32	42.1	55	8 P92501	P92501 apis mellif
9	32	42.1	61	2 Q9Z682	Q9z682 bradyrhizob
10	32	42.1	66	16 Q99RD1	Q99rd1 staphylococ
11	32	42.1	71	10 Q9LQF4	Q9lqf4 arabidopsis
12	32	42.1	78	13 Q57412	Q57412 tetraodon f
13	31	40.8	35	16 Q9KAU0	Q9kau0 bacillus ha
14	31	40.8	66	17 Q979G1	Q979g1 thermoplas
15	31	40.8	68	16 Q8YP72	Q8yp72 anabaena sp
16	31	40.8	70	4 Q96IE9	Q96ie9 homo sapien

17	31	40.8	72	2 Q9RGA4	Q9rga4 borrelia he
18	31	40.8	72	9 Q9TLJ1	Q9tlj1 lactobacill
19	31	40.8	72	17 Q8U278	Q8u278 pyrococcus
20	31	40.8	83	16 Q8EIV8	Q8eiv8 shewanella
21	31	40.8	83	16 Q8DK14	Q8dk14 synechococc
22	31	40.8	88	2 Q9RG90	Q9rg90 borrelia he
23	31	40.8	90	2 Q9RGA6	Q9rga6 borrelia he
24	31	40.8	94	2 Q9X3B0	Q9x3b0 alcaligenes
25	31	40.8	94	2 Q9X3A5	Q9x3a5 alcaligenes
26	31	40.8	96	13 Q9DFE6	Q9dff6 oncorhynch
27	31	40.8	98	16 Q9UC66	Q9uc66 staphylococ
28	31	40.8	99	16 Q8E4Q1	Q8e4q1 streptococ
29	31	40.8	100	16 Q8P2F3	Q8p2f3 streptococ
30	30.5	40.1	89	9 Q8HA80	Q8ha80 bacterioph
31	30	39.5	29	16 Q99UH5	Q99uh5 staphylococ
32	30	39.5	29	16 Q8NWX8	Q8nwx8 staphylococ
33	30	39.5	35	16 Q9PPG6	Q9ppg6 campylobact
34	30	39.5	58	17 Q8ZV98	Q8zv98 pyrobaculum
35	30	39.5	59	16 Q9KDN1	Q9kdni bacillus ha
36	30	39.5	59	16 Q8F678	Q8f678 leptospira
37	30	39.5	60	17 Q9HRS3	Q9hrs3 halobacteri
38	30	39.5	64	2 Q8GFK0	Q8gfk0 staphylococ
39	30	39.5	64	12 Q91FU2	Q91fu2 chilo iride
40	30	39.5	71	6 Q9GMS9	Q9gms9 macaca faec
41	30	39.5	72	16 Q9PDS4	Q9pds4 xyiella fae
42	30	39.5	79	16 Q9AC63	Q9ac63 staphylococ
43	30	39.5	82	12 Q98215	Q98215 molluscum c
44	30	39.5	92	10 Q8RUQ9	Q8ruc9 lolium rigi
45	30	39.5	92	10 Q8S3M1	Q8s3m1 lolium rigi

#### ALIGNMENTS

#### RESULT 1

Q8DSX0 PRELIMINARY; PRT; 68 AA.  
AC Q8DSX0;  
DT 01-MAR-2003 (Tremblrel. 23, Created)  
DT 01-MAR-2003 (Tremblrel. 23, Last sequence update)  
DT 01-MAR-2003 (Tremblrel. 23, Last annotation update)  
DE Hypothetical protein.  
GN SMU.1637C.  
OS Streptococcus mutans.  
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;  
OC Streptococcus.  
OX NCBI\_TaxID=1309;  
RN [1]

RP SEQUENCE FROM N.A.  
RC STRAIN=UAI59 / ATCC 700610 / Serotype C;  
RX MEDLINE=22295063; PubMed=12397186;  
RA Ajdic D., McShan W.M., McLaughlin R.E., Savic G., Chang J.,  
RA Carson M.B., Primeaux C., Tian R., Kenton S., Jia H., Lin S., Qian Y.,  
RA Li S., Zhu H., Najjar F., Lai H., White J., Roe B.A., Ferretti J.J.;  
RT "Genome sequence of Streptococcus mutans UAI59, a cariogenic dental  
RT pathogen.";  
RL Proc. Natl. Acad. Sci. U.S.A. 99:14434-14439(2002).  
DR EMBL: AE014994; AAN59277.1; ..  
KW Hypothetical protein; Complete proteome.

SQ SEQUENCE 68 AA; 7604 MW; EEF457B026865773 CRC64;  
Query Match 46.1%; Score 35; DB 16; Length 68;  
Best Local Similarity 53.8%; Pred. No. 1.1e+02;  
Matches 7; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

QY 4 DALNVLMMNIIS 16  
||| |  
Db 2 DALNHLAMTNLLT 14

#### RESULT 2

Q8C8H4 PRELIMINARY; PRT; 80 AA.  
ID Q8C8H4

Tue Feb 17 11:55:53 2004

AC Q8C8H4;  
 DT 01-MAR-2003 (TrEMBLrel. 23, Created)  
 DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)  
 DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)  
 DE Unknown EST (Fragment).  
 OS Mus musculus (Mouse).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 OX NCBI\_TaxID=10090;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=C57BL/6J; TISSUE=Cerebellum;  
 RX MEDLINE=22354683; PubMed=12466851;  
 RA The FANTOM Consortium,  
 RA "Analysis of the mouse transcriptome based on functional annotation of  
 RT 60,770 full-length cDNAs."  
 RL Nature 420:563-573 (2002).  
 DR EMBL; AK047089; BAC32957.1; -.  
 FT NON TER 1  
 SQ SEQUENCE 80 AA; 9655 MW; D03C342182DC4BD2 CRC64;  
 Query Match 46.1%; Score 35; DB 11; Length 80;  
 Best Local Similarity 50.0%; Pred. No. 1.3e+02;  
 Matches 6; Conservative 3; Mismatches 3; Indels 0; Gaps 0;  
 QY 2 VYDALNVLMMN 13  
 DB 18 IYNALKLFMEWN 29  
 :|||:|  
 RESULT 3  
 Q99VZ5 PRELIMINARY; PRT; 74 AA.  
 ID Q99VZ5  
 AC Q99VZ5  
 DT 01-JUN-2001 (TrEMBLrel. 17, Created)  
 DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)  
 DT 01-OCT-2002 (TrEMBLrel. 22, Last annotation update)  
 DE Hypothetical protein SAV0618 (Hypothetical protein MW0582).  
 GN SAV0618 OR SA0575 OR MW0582.  
 OS Staphylococcus aureus (strain N315), and  
 OS Staphylococcus aureus (strain N315), and  
 OS Staphylococcus aureus (strain MW2).  
 OC Bacteria; Firmicutes; Bacillales; Staphylococcus.  
 OX NCBI\_TaxID=158878, 158879, 196620;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=M50, and N315;  
 RX MEDLINE=21311952; PubMed=11418146;  
 RA Kuroda M., Ohta T., Uchiyama I., Baba T., Yuzawa H., Kobayashi I.,  
 RA Cui L., Oguchi A., Aoki K.-I., Nagai Y., Lian J.-Q., Ito T.,  
 RA Kanamori M., Matsumaru H., Maruyama A., Murakami H., Hosoyama A.,  
 RA Mizutani-Ui Y., Takahashi N.K., Sawano T., Inoue R.-I., Kaito C.,  
 RA Sekimizu K., Hirakawa H., Kuhara S., Goto S., Yabuzaki J.,  
 RA Kanehisa M., Yamaashita A., Oshima K., Furuya K., Yoshino C., Shiba T.,  
 RA Hattori M., Ogasawara N., Hayaishi H., Hiramatsu K.,  
 RA "Whole genome sequencing of methicillin-resistant Staphylococcus  
 RT aureus".  
 RL Lancet 357:1225-1240 (2001).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=MW2;  
 RX MEDLINE=22040717; PubMed=12044378;  
 RA Baba T., Takeuchi F., Kuroda M., Yuzawa H., Aoki K.-I., Oguchi A.,  
 RA Nagai Y., Iwama N., Asano K., Naimi T., Kuroda H., Cui L.,  
 RA Yamanoto K., Hiramatsu K.,  
 RA "Genome and virulence determinants of high virulence community-  
 RT acquired MRSA".  
 RL Lancet 359:1819-1827 (2002).  
 DR EMBL; AP003359; BAB56780.1; -.  
 DR EMBL; AP003131; BAB41807.1; -.  
 DR EMBL; AP004824; BAB94447.1; -.  
 KW Hypothetical protein; Complete proteome.

SQ SEQUENCE 74 AA; 8228 MW; 35AFAB2F9FF612AC CRC64;  
 Query Match 44.7%; Score 34; DB 16; Length 74;  
 Best Local Similarity 50.0%; Pred. No. 1.8e+02;  
 Matches 6; Conservative 4; Mismatches 2; Indels 0; Gaps 0;  
 QY 4 DALNVLMMNII 15  
 DB 35 DSNALVDLNI 46  
 :|||:|  
 RESULT 4  
 Q99MV9 PRELIMINARY; PRT; 80 AA.  
 ID Q99MV9  
 AC Q99MV9  
 DT 01-JUN-2001 (TrEMBLrel. 17, Created)  
 DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)  
 DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)  
 DE Testis protein TEX18.  
 GN TEX18.  
 OS Mus musculus (Mouse).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 OX NCBI\_TaxID=10090;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Testis;  
 RX MEDLINE=21175748; PubMed=11279525;  
 RA Wang P.J., McCarrey J.R., Yang F., Page D.C.;  
 RA "An abundance of X-linked genes expressed in spermatogonia".  
 RL Nat. Genet. 27:422-426 (2001).  
 DR EMBL; AF285583; AAK31962.1; -.  
 DR MGD; MGI:1890547; Tex18.  
 SQ SEQUENCE 80 AA; 9115 MW; F5CB6CE95F8EC2B1 CRC64;  
 Query Match 44.7%; Score 34; DB 11; Length 80;  
 Best Local Similarity 50.0%; Pred. No. 2e+02;  
 Matches 7; Conservative 1; Mismatches 6; Indels 0; Gaps 0;  
 QY 1 RVYDALNVLMMNI 14  
 DB 31 KYSVLNVAFACNI 44  
 :|||:|  
 RESULT 5  
 Q9HQ82 PRELIMINARY; PRT; 66 AA.  
 ID Q9HQ82  
 AC Q9HQ82;  
 DT 01-MAR-2001 (TrEMBLrel. 16, Created)  
 DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)  
 DT 01-MAR-2001 (TrEMBLrel. 16, Last annotation update)  
 DE Vng1283h.  
 GN Vng1283h.  
 OS Halobacterium sp. (strain NRC-1).  
 OC Archaea; Euryarchaeota; Halobacteria; Halobacteriales;  
 OC Halobacteriaceae; Halobacterium.  
 OX NCBI\_TaxID=64091;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC MEDLINE=20504483; PubMed=11016950;  
 RX Ng W.V., Kennedy S.P., Mahairas G.G., Berquist B., Pan M.,  
 RA Shukla H.D., Lasky S.R., Balliga N.S., Thorsson V., Sirogna J.,  
 RA Swartzell S., Weir D., Hall J., Dahl T.A., Welti R., Goo Y.A.,  
 RA Leitchauer B., Keller K., Cruz R., Danson M.J., Hough D.W.,  
 RA Maddocks D.G., Jablonski P.E., Krebs M.P., Angevine C.M., Dale H.,  
 RA Isenbarger T.A., Peck R.F., Pohlischer M., Spudich J.L., Jung K.-H.,  
 RA Alam M., Freitas T., Hou S., Daniels C.J., Dennis P.P., Omer A.D.,  
 RA Ebhardt H., Lowe T.M., Liang P., Riley M., Hood L., DasSarma S.,  
 RL "Genome sequence of Halobacterium species NRC-1".  
 RT Proc. Natl. Acad. Sci. U.S.A. 97:12176-12181 (2000).  
 DR EMBL; AE005052; AAG19634.1; -.  
 KW Complete proteome.  
 SQ SEQUENCE 66 AA; 6859 MW; B5CD27577F80E8A8 CRC64;

Query Match 43.4%; Score 33; DB 17; Length 66;  
 Best Local Similarity 50.0%; Pred. NO. 2.5e+02;  
 Matches 8; Conservative 4; Mismatches 4; Indels 0; Gaps 0;

QY 1 RYDALNVLMMNNIS 16  
 ||:||||| :||:  
 Db 20 RVFDALLVIGPWIIVA 35

RESULT 6  
 O48472 PRELIMINARY; PRT; 82 AA.  
 AC O48472;  
 DT 01-JUN-1998 (TREMBLrel. 06, Created)  
 DT 01-JUN-1998 (TREMBLrel. 06, Last sequence update)  
 DT 01-MAR-2003 (TREMBLrel. 23, Last annotation update)  
 DE Complete nucleotide sequence.  
 OS Bacteriophage SP1.  
 OC Viruses; dsDNA viruses, no RNA stage; Caudovirales; Siphoviridae;  
 OC Lambda-like viruses.  
 OC NCBI\_TaxID=10724;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Alonso J.C., Luder G., Stiege A.C., Chai S., Weise F., Trautner T.A.;  
 RT "Analysis of the complete nucleotide sequence and functional  
 organization of Bacillus subtilis bacteriophage SP1.";  
 RL Submitted (MAY-1996) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; X97918; CAA66519.1; -  
 DR InterPro; IPR006479; Holin\_SP1.  
 DR Pfam; PF04688; Phage\_holin; 1.  
 DR TIGRFAMs; TIGR01592; holin\_SP1; 1.  
 SQ SEQUENCE 82 AA; 9391 MW; 558F709AD9E0D3A0 CRC64;

Query Match 43.4%; Score 33; DB 9; Length 82;  
 Best Local Similarity 50.0%; Pred. NO. 3.1e+02;  
 Matches 6; Conservative 4; Mismatches 2; Indels 0; Gaps 0;

QY 1 RYDALNVLMM 12  
 :||||:|:  
 Db 40 QYDAVSVLFTI 51

RESULT 7  
 P92499 PRELIMINARY; PRT; 55 AA.  
 AC P92499;  
 DT 01-MAY-1997 (TREMBLrel. 03, Created)  
 DT 01-MAY-1997 (TREMBLrel. 03, Last sequence update)  
 DT 01-OCT-2002 (TREMBLrel. 22, Last annotation update)  
 DE Cytochrome oxidase subunit III (EC 1.9.3.1) (Cytochrome c oxidase polypeptide III) (Fragment).  
 GN COIII  
 OS Apis mellifera (Honeybee).  
 OC Mitochondrion.  
 OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;  
 OC Neoptera; Endopterygota; Hymenoptera; Apocrita; Aculeata; Apoidea;  
 OC Apidae;  
 OC NCBI\_TaxID=7460;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Koulianos S., Crozier R.H.;  
 RT "Mitochondrial sequence characterisation of Australian commercial and feral honeybee strains (Hymenoptera: Apidae: Apis mellifera Linnaeus), in the context of the species worldwide.";  
 RL J. Aust. Entomol. Soc. 0:0-0(1997).  
 CC -1- FUNCTION: SUBUNITS I, II AND III FORM THE FUNCTIONAL CORE OF THE ENZYME COMPLEX (BY SIMILARITY).  
 CC -1- CATALYTIC ACTIVITY: 4 FERROCYTOCHROME C + O(2) = 4 FERRICYTOCHROME C + 2 H(2)O.  
 CC -1- SIMILARITY: BELONGS TO THE CYTOCHROME C OXIDASE SUBUNIT 3 FAMILY.  
 DR EMBL; U72269; AAB41167.1; -  
 DR InterPro; IPR000298; CytC\_oxdse\_III.

DR Pfam; PF00510; COX3; 1.  
 DR ProDom; PD000382; CytC\_oxdse\_III; 1.  
 DR PROSITE; PS0253; COX3; 1.  
 KW Oxidoreductase; Transmembrane; Mitochondrion.  
 FT NON\_TER 55  
 SQ SEQUENCE 55 AA; 6506 MW; 650A864ED7C2FC0C CRC64;

Query Match 42.1%; Score 32; DB 8; Length 55;  
 Best Local Similarity 26.7%; Pred. NO. 3.1e+02;  
 Matches 4; Conservative 7; Mismatches 4; Indels 0; Gaps 0;

QY 2 VYDALNVLMMNNIS 16  
 :|:::|:|:  
 Db 34 IYSSISMFILNFN 48

RESULT 8  
 P92501 PRELIMINARY; PRT; 55 AA.  
 AC P92501;  
 DT 01-MAY-1997 (TREMBLrel. 03, Created)  
 DT 01-MAY-1997 (TREMBLrel. 03, Last sequence update)  
 DT 01-OCT-2002 (TREMBLrel. 22, Last annotation update)  
 DE Cytochrome oxidase subunit III (EC 1.9.3.1) (Cytochrome c oxidase polypeptide III) (Fragment).  
 GN COIII.  
 OS Apis mellifera (Honeybee).  
 OC Mitochondrion.  
 OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;  
 OC Neoptera; Endopterygota; Hymenoptera; Apocrita; Aculeata; Apoidea;  
 OC Apidae; Apis.  
 OC NCBI\_TaxID=7460;  
 RN [1]  
 RP SEQUENCE FROM N.A.

RA Koulianos S., Crozier R.H.;  
 RT "Mitochondrial sequence characterisation of Australian commercial and feral honeybee strains (Hymenoptera: Apidae: Apis mellifera Linnaeus), in the context of the species worldwide.";  
 RL J. Aust. Entomol. Soc. 0:0-0(1997).  
 CC -1- FUNCTION: SUBUNITS I, II AND III FORM THE FUNCTIONAL CORE OF THE ENZYME COMPLEX (BY SIMILARITY).  
 CC -1- CATALYTIC ACTIVITY: 4 FERROCYTOCHROME C + O(2) = 4 FERRICYTOCHROME C + 2 H(2)O.  
 CC -1- SIMILARITY: BELONGS TO THE CYTOCHROME C OXIDASE SUBUNIT 3 FAMILY.  
 DR EMBL; U72271; AAB41169.1; -  
 DR InterPro; IPR000298; CytC\_oxdse\_III.  
 DR Pfam; PF00510; COX3; 1.  
 DR ProDom; PD000382; CytC\_oxdse\_III; 1.  
 DR PROSITE; PS0253; COX3; 1.  
 KW Oxidoreductase; Transmembrane; Mitochondrion.  
 FT NON\_TER 55  
 SQ SEQUENCE 55 AA; 6488 MW; 7E10364ED7C2FC0D CRC64;

Query Match 42.1%; Score 32; DB 8; Length 55;  
 Best Local Similarity 26.7%; Pred. NO. 3.1e+02;  
 Matches 4; Conservative 7; Mismatches 4; Indels 0; Gaps 0;

QY 2 VYDALNVLMMNNIS 16  
 :|:::|:|:  
 Db 34 IYSSISMFILNFN 48

RESULT 9  
 Q92682 PRELIMINARY; PRT; 61 AA.  
 AC Q92682;  
 DT 01-MAY-1999 (TREMBLrel. 10, Created)  
 DT 01-MAY-1999 (TREMBLrel. 10, Last sequence update)  
 DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)  
 DE Exonuclease I (Fragment).  
 OS Bradyrhizobium japonicum.  
 OC Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;  
 OC Bradyrhizobiaceae; Bradyrhizobium.

Tue Feb 17 11:55:53 2004

us-09-900-147-5.isrpt

```

OX NCBI_TaxID=375;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=USDA110;
RA Mayer R.M., Mathis J.N., McMillin D.E.;
RT "Analysis of a DNA fragment present in Bradyrhizobium japonicum 110
RT Fix+ derivatives but missing in a Fix- derivative.";
RL Submitted (OCT-1998) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF101073; AAD17890.1; -
FT NON TER 61
SQ SEQUENCE 61 AA; 6729 MW; FF29A9A1891C986A CRC64;

Query Match 42.1%; Score 32; DB 2; Length 61;
Best Local Similarity 46.2%; Pred. NO. 3.5e+02;
Matches 6; Conservative 5; Mismatches 2; Indels 0; Gaps 0;

Qy 4 DALANVLMAMNIIIS 16
Db 25 DALSLVMAADAVS 37

RESULT 10
Q99RD1 PRELIMINARY; PRT; 66 AA.
AC Q99RD1;
DT 01-JUN-2001 (TrEMBLrel. 17, Created)
DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT 01-OCT-2002 (TrEMBLrel. 22, Last annotation update)
DE Hypothetical protein SAV2504 (Hypothetical protein MW2422).
GN SAV2504 OR SA2292 OR MW2422
OS Staphylococcus aureus (strain Mu50 / ATCC 700699),
OS Staphylococcus aureus (strain N315), and
OS Staphylococcus aureus (strain MW2).
OC Bacteria; Firmicutes; Bacillales; Staphylococcus.
OX NCBI_TaxID=158878, 158879, 196620;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Mu50, and N315;
RX MEDLINE=22040717; PubMed=12044378;
RA Kuroda M., Ohta T., Uchiyama I., Baba T., Yuzawa H., Kobayashi I.,
RA Cui L., Oguchi A., Aoki K.-I., Nagai Y., Lian J.-Q., Ito T.,
RA Kanamori M., Matsumaru H., Maruyama A., Murakami H., Hosoyama A.,
RA Mizutani-Ui Y., Takahashi N.K., Sawano T., Inoue R.-I., Kaito C.,
RA Sekimizu K., Hirakawa H., Kuhara S., Goto S., Yabuzaki J.,
RA Kanehisa M., Masahita A., Oshima K., Furuya K., Yoshino C., Shiba T.,
RA Hattori M., Ogasawara N., Hayashi H., Hiramatsu K.;
RT "Whole genome sequencing of methicillin-resistant Staphylococcus
RT aureus.";
RL Lancet 357:1225-1240(2001).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=MW2;
RX MEDLINE=22040717; PubMed=12044378;
RA Baba T., Takeuchi F., Kuroda M., Yuzawa H., Aoki K.-I., Oguchi A.,
RA Nagai Y., Iwama N., Asano K., Naimi T., Kuroda H., Cui L.,
RA Yamamoto K., Hiramatsu K.;
RT "Genome and virulence determinants of high virulence community-
RT acquired MRSA.";
RL Lancet 359:1819-1827(2002).
DR EMBL; AP003365; BAB58666.1; -
DR EMBL; AP003137; BAB43595.1; -
DR EMBL; AP004830; BAB96287.1; -
SQ SEQUENCE 66 AA; 8016 MW; 8B6B4B3627F791C CRC64;

Query Match 42.1%; Score 32; DB 16; Length 66;
Best Local Similarity 66.7%; Pred. NO. 3.7e+02;
Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Qy 1 RVDALNVL 9
Db 47 RYVDLNTI 55

RESULT 11
Q9LQF4 PRELIMINARY; PRT; 71 AA.
AC Q9LQF4;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-OCT-2000 (TrEMBLrel. 15, Last annotation update)
DE F1504.31.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eurosids II; Brassicales; Brassicaceae; Arabidopsi.
OX NCBI_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.
RA Ecker J.R.;
RL Submitted (JUN-1999) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RA Ecker J.R.;
RL Submitted (OCT-1999) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RA Cheuk R., Shinn P., Brooks S., Buehler E., Chao Q., Johnson-Hopson C.,
RA Khan S., Kim C., Altafi H., Bei B., Chin C., Chlou J., Choi E.,
RA Conn L., Conway A., Gonzalez A., Hansen N., Howing B., Koo T., Lam B.,
RA Lee J., Lenz C., Li J., Liu A., Liu J., Liu S., Mukharsky N.,
RA Nguyen M., Palm C., Pham P., Sakano H., Schwartz J., Southwick A.,
RA Thaveri A., Toriumi M., Vaysberg M., Yu G., Davis R., Federspiel N.,
RA Theologis A., Ecker J.;
RL Submitted (JUN-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AC007887; AAF79373.1; -
SQ SEQUENCE 71 AA; 8569 MW; DA361597E8A05E7E CRC64;

Query Match 42.1%; Score 32; DB 10; Length 71;
Best Local Similarity 33.3%; Pred. NO. 4e+02;
Matches 5; Conservative 5; Mismatches 5; Indels 0; Gaps 0;

Qy 1 RYDANVLMAMNII 15
Db 35 RYDMMNIISSQLVVI 49

RESULT 12
Q57412 PRELIMINARY; PRT; 78 AA.
ID Q57412
AC Q57412;
DT 01-JUN-1998 (TrEMBLrel. 06, Created)
DT 01-JUN-1998 (TrEMBLrel. 06, Last sequence update)
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE Guanine nucleotide dissociation stimulator (Fragment).
GN GNDS
OS Tetraodon fluviatilis (Puffer fish).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
OC Acanthomorpha; Acanthopterygii; Percomorpha; Tetraodontiformes;
OC Tetraodontidae; Tetraodontidae; Tetraodon.
OX NCBI_TaxID=47145;
RN [1]
RP SEQUENCE FROM N.A.
RA Yao C.W., Huang C.J.;
RT "partial genomic sequence of GNDS gene of puffer fish (Tetraodon
RT fluviatilis).";
RL Submitted (JUN-1997) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF007814; AAB94740.1; -
DR HSP; Q12967; 2RCF
DR InterPro; IPR000463; Fatty_acid_BP.
DR InterPro; IPR000159; RA_domain.
DR Pfam; PF00788; RA; 1.
DR PROSITE; PS00214; FAFP; 1.
DR NON_TER 1
SQ SEQUENCE 78 AA; 9109 MW; 5EDA65DB7C297A16 CRC64;

```

Query Match 42.1%; Score 32; DB 13; Length 78;  
 Best Local Similarity 53.8%; Pred. No. 4.4e+02;  
 Matches 7; Conservative 1; Mismatches 5; Indels 0; Gaps 0;

QY 1 RYDNLNVLMMN 13  
 DB 25 RIPDNANVFYMN 37

RESULT 13  
 Q9KAU0 PRELIMINARY; PRT; 35 AA.  
 ID Q9KAU0  
 AC Q9KAU0  
 DT 01-OCT-2000 (TrEMBLrel. 15, Created)  
 DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)  
 DT 01-MAR-2002 (TrEMBLrel. 20, Last annotation update)  
 DE Hypothetical protein BH2196.  
 GN BH2196.  
 OS Bacillus halodurans.  
 OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.  
 OX NCBI\_TaxID=86665;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=C-125 / JCM 9153;  
 RX MEDLINE=20512582; PubMed=11058132;  
 RA Takami H., Nakasone K., Takaki Y., Maeno G., Sasaki R., Masui N.,  
 RA Fuji F., Hirama C., Nakamura Y., Ogasawara N., Kuhara S.,  
 RA Horikoshi K.;  
 RT "Complete genome sequence of the alkaliphilic bacterium Bacillus  
 RT halodurans and genomic sequence comparison with Bacillus subtilis.";  
 RL Nucleic Acids Res. 28:4317-4331(2000).  
 DR EMBL; AP001514; BAB05915.1; -;  
 KW Hypothetical protein; Complete proteome.  
 SQ SEQUENCE 35 AA; 3960 MW; 9f851af909db2388 CRC64;

Query Match 40.8%; Score 31; DB 16; Length 35;  
 Best Local Similarity 35.7%; Pred. No. 3e+02;  
 Matches 5; Conservative 6; Mismatches 3; Indels 0; Gaps 0;

QY 2 VYDALNVLMMNII 15  
 DB 3 VYELSIIVQVNSI 16

RESULT 14  
 Q979G1 PRELIMINARY; PRT; 66 AA.  
 ID Q979G1  
 AC Q979G1  
 DT 01-OCT-2001 (TrEMBLrel. 18, Created)  
 DT 01-OCT-2001 (TrEMBLrel. 18, Last sequence update)  
 DT 01-MAR-2002 (TrEMBLrel. 20, Last annotation update)  
 DE Hypothetical protein TV1200.  
 GN TV1200 OR TVG1229743.  
 OS Thermoplasma volcanium.  
 OC Archaea; Euryarchaeota; Thermoplasmatia; Thermoplasmatiales;  
 OC Thermoplasmataceae; Thermoplasma.  
 OX NCBI\_TaxID=50339;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=GSS1 / DSM 4299 / JCM 9571;  
 RX MEDLINE=20570466; PubMed=11121031;  
 RA Kawashima T., Anano N., Koike H., Makino S.-I., Higuchi S.,  
 RA Kawashima-Ohya Y., Watanabe K., Yamazaki M., Kanehori K., Kawamoto T.,  
 RA Nunoshiba T., Yamamoto Y., Aramaki H., Makino K., Suzuki M.;  
 RT "Archaeal adaptation to higher temperatures revealed by genomic  
 RT sequence of Thermoplasma volcanium.";  
 RL Proc. Natl. Acad. Sci. U.S.A. 97:14257-14262(2000).  
 DR EMBL; AP000995; BAB60342.1; -;  
 KW Hypothetical protein; Complete proteome.  
 SQ SEQUENCE 66 AA; 7786 MW; 7237a16a61e309fb CRC64;

Query Match 40.8%; Score 31; DB 17; Length 66;

Best Local Similarity 54.5%; Pred. No. 5.7e+02;  
 Matches 6; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 6 LNVLMAMNIIIS 16  
 DB 50 LNIIVDKNIIIS 60

RESULT 15  
 Q8YP72 PRELIMINARY; PRT; 68 AA.  
 ID Q8YP72  
 AC Q8YP72  
 DT 01-MAR-2002 (TrEMBLrel. 20, Created)  
 DT 01-MAR-2002 (TrEMBLrel. 20, Last sequence update)  
 DT 01-MAR-2002 (TrEMBLrel. 20, Last annotation update)  
 DE Hypothetical protein Asl4328.  
 GN Asl4328.  
 OS Anabaena sp. (strain PCC 7120).  
 OC Bacteria; Cyanobacteria; Nostocales; Nostocaceae; Nostoc.  
 OX NCBI\_TaxID=103690;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC MEDLINE=21595285; PubMed=11759840;  
 RX Kaneko T., Nakamura Y., Wolk C.P., Kuritz T., Sasamoto S.,  
 RA Watanabe A., Iriguchi M., Ishikawa A., Kawashima K., Kimura T.,  
 RA Kishida Y., Kohara M., Matsumoto M., Matsuno A., Muraki A.,  
 RA Nakazaki N., Shimo S., Sugimoto M., Takazawa M., Yamada M.,  
 RA Yasuda M., Tabata S.;  
 RT "Complete genomic sequence of the filamentous nitrogen-fixing  
 RT cyanobacterium Anabaena sp. strain PCC 7120.";  
 RL DNA Res. 8:205-213(2001).  
 DR EMBL; AP003596; BAB76027.1; -;  
 KW Hypothetical protein; Complete proteome.  
 SQ SEQUENCE 68 AA; 7648 MW; 8876d5d9ffcc4b14 CRC64;

Query Match 40.8%; Score 31; DB 16; Length 68;  
 Best Local Similarity 66.7%; Pred. No. 5.8e+02;  
 Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 3 YDALNVLMA 11  
 DB 43 YEALLELMA 51

RESULT 16  
 Q96IE9 PRELIMINARY; PRT; 70 AA.  
 ID Q96IE9  
 AC Q96IE9  
 DT 01-DEC-2001 (TrEMBLrel. 19, Created)  
 DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)  
 DT 01-OCT-2002 (TrEMBLrel. 22, Last annotation update)  
 DE Hypothetical protein.  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 OX NCBI\_TaxID=9606;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Placenta;  
 RA Strausberg R.;  
 RL Submitted (May-2001) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; BC007583; AAH07583.1; -;  
 DR InterPro; IPR001014; Ribosomal\_L23.  
 DR Pfam; PF00276; Ribosomal\_L23; 1.  
 DR ProDom; PD001141; Ribosomal\_L23; 1.  
 DR PROSITE; PS00050; RIBOSOMAL\_L23; 1.  
 KW Hypothetical protein.  
 SQ SEQUENCE 70 AA; 7923 MW; AC1C466548f343c1 CRC64;

Query Match 40.8%; Score 31; DB 4; Length 70;  
 Best Local Similarity 53.8%; Pred. No. 6e+02;  
 Matches 7; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

Tue Feb 17 11:55:53 2004

us-09-900-147-5.rapt

```

OY 3 YDALNVLMMNNII 15
DB 58 YDALDVANKIGII 70

RESULT 17
O9RGGA PRELIMINARY; PRT; 72 AA.
AC O9RGGA4;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE BdrB1-.
DE BDRB1-.
GN Borrelia hermsii.
OS Bacteria; Spirochaetes; Spirochaetales; Spirochaetaceae; Borrelia.
OC NCBI_TaxID=140;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Man;
RA "Carlyon J.A., Roberts D.M., Theisen M., Sadler C., Marconi R.T.;
RT "Molecular and immunological analyses of the B. turicatae Bdr protein
RT family: a polymorphic, linear plasmid carried, paralogous gene
RT family."
RL Infect. Immun. 0:0-0(2000).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=Man;
RX MEDLINE=20112991; PubMed=10644495;
RA "Carlyon J.A., Roberts D.M., Marconi R.T.;
RT "Evolutionary and molecular analyses of the Borrelia bdr super gene
RT family: Delineation of distinct sub-families and demonstration of the
RT genus wide conservation of putative functional domains, structural
RT properties and repeat motifs."
RL Microb. Pathog. 28:89-105(2000).
DR EMBL; AFI43463; AAF19121.1; -.
DR InterPro; IPR004874; Borrelia_rep.
DR Pfam; PF03183; Borrelia_rep; 4.
SQ SEQUENCE 72 AA; 8090 MW; 4C20FFCC4487F35A CRC64;

Query Match 40.8%; Score 31; DB 2; Length 72;
Best Local Similarity 25.0%; Pred. No. 6.2e+02;
Matches 4; Conservative 9; Mismatches 3; Indels 0; Gaps 0;

OY 1 RYDALNVLMMNNII 16
DB 34 RLFSTVSILLRSLIS 49

RESULT 18
O9T1J1 PRELIMINARY; PRT; 72 AA.
AC O9T1J1;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE Hypothetical 8.0 kDa protein.
OS Lactobacillus bacteriophage phi adh.
OC Viruses; dsDNA viruses, no RNA stage; Caudovirales; Siphoviridae.
OX NCBI_TaxID=12417;
RN [1]
RP SEQUENCE FROM N.A.
RA Altermann E.;
RL Submitted (DEC-1998) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=99384014; PubMed=10452953;
RA Altermann E., Klein J., Heinrich B.;
RT "Primary structure and features of the genome of the Lactobacillus
RT gasserii temperate bacteriophage phi-adh.";
RL Gene 236:333-346(1999).
RN [3]
RP SEQUENCE FROM N.A.

MEDLINE=95138034; PubMed=7836307;
RX Heinrich B., Binshofer B., Blaesi U.;
RT "Primary structure and functional analysis of the lysis genes of
RT Lactobacillus gasserii bacteriophage phi-adh.";
RL J. Bacteriol. 177:723-732(1995).
RN [4]
RP SEQUENCE FROM N.A.
RX MEDLINE=93231538; PubMed=8472961;
RA Fremaux C., De Antoni G., Raya R., Klaenhammer T.;
RT "Genetic organization and sequence of the region encoding integrative
RT functions from Lactobacillus gasserii temperate bacteriophage phi-
RT adh.";
RL Gene 126:61-66(1993).
RN [5]
RP SEQUENCE FROM N.A.
RA Engel G., Altermann E., Klein J., Heinrich B.;
RT "Structure of a genome region of the Lactobacillus gasserii temperate
RT phage phi adh covering a repressor gene and cognate promoters.";
RL Gene 210:67-70(1998).
DR EMBL; AUI31519; CAB52486.1; -.
DR InterPro; IPR000209; Peptidase_S8.
DR PROSITE; PS00136; SUBTILASE_ASF; 1.
KW Hypothetical protein.
SQ SEQUENCE 72 AA; 8045 MW; 773F8886575C889A CRC64;

Query Match 40.8%; Score 31; DB 9; Length 72;
Best Local Similarity 58.3%; Pred. NO. 6.2e+02;
Matches 7; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

OY 2 YVDALNVLMMNN 13
DB 6 VKDALKELVN 17

RESULT 19
O8U278 PRELIMINARY; PRT; 72 AA.
ID O8U278;
AC O8U278;
DT 01-JUN-2002 (TrEMBLrel. 21, Created)
DT 01-JUN-2002 (TrEMBLrel. 21, Last sequence update)
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE Hypothetical protein PF0961.
GN PF0961.
OS Pyrococcus furiosus.
OC Archaea; Euryarchaeota; Thermococci; Thermococcales; Thermococcaceae;
OC Pyrococcus.
OX NCBI_TaxID=2261;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Vc1 / DSM 3638 / ATCC 43587 / JCM 8422;
RA Weiss R.B., Dunn D.M., Robb F.T., Brown J.R.;
RT "The complete sequence of the Pyrococcus furiosus genome."
RL Submitted (FEB-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AE010209; AAL81085.1; -.
KW Hypothetical protein; Complete proteome.
SQ SEQUENCE 72 AA; 8246 MW; CA74EC8FB8371720 CRC64;

Query Match 40.8%; Score 31; DB 17; Length 72;
Best Local Similarity 28.6%; Pred. NO. 6.2e+02;
Matches 4; Conservative 9; Mismatches 1; Indels 0; Gaps 0;

OY 1 RYDALNVLMMNNI 14
DB 32 RLFNSINILIALSV 45

RESULT 20
O8EIV8 PRELIMINARY; PRT; 83 AA.
ID O8EIV8;
AC O8EIV8;
DT 01-MAR-2003 (TrEMBLrel. 23, Created)
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)

```

```

DE Conserved hypothetical protein.
GN SO0721.
OS Shewanella oneidensis.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Alteromonadales;
OC Alteromonadaceae; Shewanella.
OX NCBI_TaxID=70863;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=MR-1;
RX MEDLINE=22297686; PubMed=12368813;
RA Heidelberg J.F., Paulsen I.T., Nelson K.E., Gaidos E.J., Nelson W.C.,
RA Read T.D., Eisen J.A., Seshadri R., Ward N., Methe B., Clayton R.A.,
RA Meyer T., Tsapin A., Scott J., Beanan M., Brinkac L., Daugherty S.,
RA Deboy R.T., Dodson R.J., Durkin A.S., Haft D.H., Kolonay J.F.,
RA Madupu R., Peterson J.D., Umayam L.A., White O., Wolf A.M.,
RA Vamathevan J., Weidman J., Impraim M., Lee K., Berry K., Lee C.,
RA Mueller J., Khouri H., Gill J., Uterback T.R., McDonald L.A.,
RA Feldblyum T.V., Smith H.O., Venter J.C., Nealon K.H., Fraser C.M.;
RT "Genome sequence of the dissimilatory metal ion-reducing bacterium
RT Shewanella oneidensis.";
RL Nat. Biotechnol. 20:1118-1123(2002).
DR EMBL; AB015517; AAN53799.1; -.
KW TIGR; SO0721; -.
KW Hypothetical protein; Complete proteome.
SQ SEQUENCE 83 AA; 9075 MW; AC5D08F38ACB345C CRC64;

Query Match 40.8%; Score 31; DB 16; Length 83;
Best Local Similarity 56.2%; Pred. No. 7.1e+02;
Matches 9; Conservative 2; Mismatches 3; Indels 2; Gaps 1;

Qy 2 VYDAL--NVLMMNII 15
Db 21 LFOALTDNPLMANGII 36

RESULT 21
Q8DKI4
ID Q8DKI4 PRELIMINARY; PRT; 83 AA.
AC Q8DKI4;
DT 01-MAR-2003 (TRENBLrel. 23, Created)
DT 01-MAR-2003 (TRENBLrel. 23, Last sequence update)
DE Tsl0875 protein.
GN Tsl0875.
OS Synechococcus elongatus (Thermosynechococcus elongatus).
OC Bacteria; Cyanobacteria; Chroococcales; Synechococcus.
OX NCBI_TaxID=32046;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=BP-1;
RX MEDLINE=22251144; PubMed=12240834;
RA Nakamura Y., Kaneko T., Sato S., Ikeuchi M., Katoh H., Sasamoto S.,
RA Watanabe A., Iriguchi M., Kawashina K., Kimura T., Kishida Y.,
RA Kiyokawa C., Kohara M., Matsunoto M., Matsuno A., Nakazaki N.,
RA Shimo S., Sugimoto M., Takeuchi C., Yamada M., Tabata S.;
RT "Complete genome structure of the thermophilic cyanobacterium
RT Thermosynechococcus elongatus BP-1.";
RL DNA Res. 9:123-130(2002).
DR EMBL; AP005371; BAC08427.1; -.
KW Complete proteome.
SQ SEQUENCE 83 AA; 9297 MW; BE8A0246CDA7E1DD CRC64;

Query Match 40.8%; Score 31; DB 16; Length 83;
Best Local Similarity 58.3%; Pred. No. 7.1e+02;
Matches 7; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

Qy 2 VYDALNVLMMN 13
Db 53 VYSSKLDMASN 64

RESULT 22
Q9RG90
ID Q9RG90 PRELIMINARY; PRT; 88 AA.
AC Q9RG90;
DT 01-MAY-2000 (TRENBLrel. 13, Created)
DT 01-MAY-2000 (TRENBLrel. 13, Last sequence update)
DT 01-JUN-2002 (TRENBLrel. 21, Last annotation update)
DE BdrC4-.
GN BDRC4-.
OS Borrelia hermsii.
OC Bacteria; Spirochaetes; Spirochaetales; Spirochaetaceae; Borrelia.
OX NCBI_TaxID=140;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Man;
RA Carlyon J.A., Roberts D.M., Theisen M., Sadler C., Marconi R.T.;
RT "Molecular and immunological analyses of the B. turicatae Bdr protein
RT family: a polymorphic, linear plasmid carried, paralogous gene
RT family.";
RL Infect. Immun. 0:0-0(2000).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=Man;
RX MEDLINE=20112991; PubMed=10644495;
RA Carlyon J.A., Roberts D.M., Marconi R.T.;
RT "Evolutionary and molecular analyses of the Borrelia bdr super gene
RT family: Delineation of distinct sub-families and demonstration of the
RT genus wide conservation of putative functional domains, structural
RT properties and repeat motifs.";
RL Microb. Pathog. 28:89-105(2000).
DR EMBL; AF143471; AAF19137.1; -.
DR InterPro; IPR004874; Borrelia_rep.
DR Pfam; PF03183; Borrelia_rep; 4.
SQ SEQUENCE 88 AA; 9915 MW; B41EDF014D35B0A4 CRC64;

Query Match 40.8%; Score 31; DB 2; Length 88;
Best Local Similarity 25.0%; Pred. No. 7.5e+02;
Matches 4; Conservative 9; Mismatches 3; Indels 0; Gaps 0;

Qy 1 RYVDALNVLMMNIIIS 16
Db 32 RLFTSVSILLRSLIS 47

RESULT 23
Q9RGA6
ID Q9RGA6 PRELIMINARY; PRT; 90 AA.
AC Q9RGA6;
DT 01-MAY-2000 (TRENBLrel. 13, Created)
DT 01-MAY-2000 (TRENBLrel. 13, Last sequence update)
DT 01-JUN-2002 (TRENBLrel. 21, Last annotation update)
DE BdrC4-.
GN BDRC4-.
OS Borrelia hermsii.
OC Bacteria; Spirochaetes; Spirochaetales; Spirochaetaceae; Borrelia.
OX NCBI_TaxID=140;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=HS1;
RA Carlyon J.A., Roberts D.M., Theisen M., Sadler C., Marconi R.T.;
RT "Molecular and immunological analyses of the B. turicatae Bdr protein
RT family: a polymorphic, linear plasmid carried, paralogous gene
RT family.";
RL Infect. Immun. 0:0-0(2000).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=HS1;
RX MEDLINE=20112991; PubMed=10644495;
RA Carlyon J.A., Roberts D.M., Marconi R.T.;
RT "Evolutionary and molecular analyses of the Borrelia bdr super gene
RT family: Delineation of distinct sub-families and demonstration of the
RT genus wide conservation of putative functional domains, structural
RT properties and repeat motifs.";
RL Microb. Pathog. 28:89-105(2000).
DR EMBL; AF143462; AAF19119.1; -.

```

Tue Feb 17 11:55:53 2004

```

DR InterPro: IPR004874; Borrelia rep.
DR Pfam: PF03183; Borrelia rep; 5
SQ SEQUENCE 90 AA; 10131 MW; C92CFFPF19025846B CRC64;

Query Match 40.8%; Score 31; DB 2; Length 90;
Best Local Similarity 25.0%; Pred. No. 7.e+02;
Matches 4; Conservative 9; Mismatches 3; Indels 0; Gaps 0;

Qy 1 RYDVALNVLMMNII 16
Db 34 RLFTVSILLRSLS 49

RESULT 24
ID Q9X3B0 PRELIMINARY; PRT; 94 AA.
AC Q9X3B0;
DT 01-NOV-1999 (TrEMBLrel. 12, Created)
DT 01-NOV-1999 (TrEMBLrel. 12, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE 2,4-dichlorophenol hydroxylase (Fragment).
GN TFDB.
OS Alcaligenes sp. JD12.
OC Bacteria; Proteobacteria; Betaproteobacteria; Burkholderiales;
OC Alcaligenaceae; Alcaligenes.
OX NCBI_TaxID=87548;
RN [1]
RP SEQUENCE FROM N.A.

RC STRAIN=JD12;
RA Vallaeys T., Fulthorpe R.R., Lau P., Wright A.D.;
RT "Comparison of ttdB sequences from bacterial isolates degrading 2,4-
RT dichlorophenoxyacetic acid and 2,4-dichlorophenol.";
RL Abstr. Gen. Meet. Am. Soc. Microbiol. 97:Q403-Q403 (1997).
RN [2]
RP SEQUENCE FROM N.A.

RC STRAIN=JD12;
RA Vallaeys T., Courde L., McGowan C., Wright A.D., Fulthorpe R.R.;
RT "Phylogenetic analyses indicate independent recruitment of diverse
RT gene cassettes during assemblage of the 2,4-D catabolic pathway.";
RL FEMS Microbiol. Ecol. 0:0-0 (1999).
DR EMBL; AF068274; AAD24417.1; -.
DR InterPro: IPR002202; HMG-CoA red.
DR InterPro: IPR002938; MoxY FAD binding.
DR Pfam; PF01494; FAD binding_3; 1.
DR PROSITE; PS00066; HMG_COA_REDUCTASE_1; 1.
FT NON_TER 1
FT NON_TER 94
SQ SEQUENCE 94 AA; 10499 MW; 9EB81C9689FA30E3 CRC64;

Query Match 40.8%; Score 31; DB 2; Length 94;
Best Local Similarity 53.3%; Pred. No. 8.e+02;
Matches 8; Conservative 1; Mismatches 6; Indels 0; Gaps 0;

Qy 1 RYDVALNVLMMNII 15
Db 11 RLYAAPNDLMGENTI 25

RESULT 26
ID Q9DFP6 PRELIMINARY; PRT; 96 AA.
AC Q9DFP6;
DT 01-MAR-2001 (TrEMBLrel. 16, Created)
DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT 01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
DE 60S ribosomal protein (Fragment).
OS Oncorhynchus mykiss (Rainbow trout) (Salmo gairdneri).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Euteleostei;
OC Protacanthopterygii; Salmoniformes; Salmonidae; Oncorhynchus.
OX NCBI_TaxID=8022;
RN [1]
RP SEQUENCE FROM N.A.
RA Bayne C.J., Gerwick L., Fujiki K., Nakao M., Yano T.;
RT "Immune-relevant (including acute phase) genes identified in the
RT livers of rainbow trout, Oncorhynchus mykiss, by means of suppression
RT subtractive hybridization.";
RL Submitted (JUN-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF281334; AAG30009.1; -.
DR InterPro: IPR001014; Ribosomal_L23.
DR Pfam; PF00276; Ribosomal_L23; 1.
DR ProDom; PD001141; Ribosomal_L23; 1.
DR PROSITE; PS00050; RIBOSOMAL_L23; 1.
KW Ribosomal protein.
FT NON_TER 1
FT NON_TER 94
SQ SEQUENCE 96 AA; 10915 MW; 4C01DE3F3E37F906 CRC64;

Query Match 40.8%; Score 31; DB 13; Length 96;
Best Local Similarity 53.8%; Pred. No. 8.1e+02;
Matches 7; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

Qy 3 YDALNVLMMNII 15
Db 84 YDALDVANKIGII 96

RESULT 27
ID Q99UC6 PRELIMINARY; PRT; 98 AA.
AC Q99UC6;
DT 01-JUN-2001 (TrEMBLrel. 17, Created)

```



01-JUN-2001 (TRENBLrel. 17, Last sequence update)  
 01-OCT-2002 (TRENBLrel. 22, Last annotation update)  
 DE Hypothetical protein SAV1352 (Hypothetical protein MW1239).  
 GN SAV1352 OR SAI186 OR MW1239.  
 OS Staphylococcus aureus (strain Mu50 / ATCC 700699),  
 OS Staphylococcus aureus (strain N315), and  
 OS Staphylococcus aureus (strain MW2).  
 OC Bacteria; Firmicutes; Bacillales; Staphylococcus.  
 OX NCBI\_TaxID=158878, 158879, 196620;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=Mu50. and N315;  
 RX MEDLINE=21311952; PubMed=11418146;  
 RA Kuroda M., Ohta T., Uchiyama I., Baba T., Yuzawa H., Kobayashi I.,  
 RA Cui L., Oguchi A., Aoki K.-I., Nagai Y., Lian J.-Q., Ito Y.,  
 RA Kanamori M., Matsumaru H., Maruyama A., Murakami H., Hosoyama A.,  
 RA Mitutani U. Y., Takahashi N. K., Sawano T., Inoue R.-I., Kaito C.,  
 RA Sekimizu K., Hirakawa H., Kuhara S., Goto S., Yabuzaki J.,  
 RA Kanehisa M., Yanashita A., Oshima K., Furuya K., Yoshino C., Shiba T.,  
 RA Hattori M., Ogasawara N., Hayashi H., Hiramatsu K.;  
 RT "Whole genome sequencing of methicillin-resistant Staphylococcus  
 aureus.";  
 RL Lancet 357:1225-1240(2001).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=MW2;  
 RX MEDLINE=22040717; PubMed=12044378;  
 RA Baba T., Takeuchi F., Kuroda M., Yuzawa H., Aoki K.-I., Oguchi A.,  
 RA Nagai Y., Iwano N., Asano K., Naimi T., Kuroda H., Cui L.,  
 RA Yamamoto K., Hiramatsu K.;  
 RT "Genome and virulence determinants of high virulence community-  
 acquired MRSA.";  
 RL Lancet 359:1819-1827(2002).  
 DR EMBL; AP003362; BAB57514.1; --  
 DR EMBL; AP003133; BAB2444.1; --  
 DR EMBL; AP004826; BAB95104.1; --  
 KW Hypothetical protein; Complete proteome.  
 SQ SEQUENCE 98 AA; 11496 MW; 216697709B904141 CRC64;  
 Query Match 40.8%; Score 31; DB 16; Length 98;  
 Best Local Similarity 66.7%; Pred. No. 8.3e+02;  
 Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;  
 Qy 3 YDNLNVLMA 11  
 Db 62 YDNLNVVA 70  
 RESULT 28  
 Q8E4Q1 PRELIMINARY; PRT; 99 AA.  
 ID Q8E4Q1  
 AC Q8E4Q1;  
 DT 01-MAR-2003 (TRENBLrel. 23, Created)  
 DT 01-MAR-2003 (TRENBLrel. 23, Last sequence update)  
 DT 01-MAR-2003 (TRENBLrel. 23, Last annotation update)  
 DE Hypothetical protein.  
 GN GBS1350.  
 OS Streptococcus agalactiae (serotype III).  
 OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;  
 OX NCBI\_TaxID=216495;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=NEM316 / Serotype III;  
 RX MEDLINE=22242508; PubMed=12354221;  
 RA Glaser P., Ruenloek C., Buchrieser C., Chevalier F., Frangoul L.,  
 RA Msadek T., Zouine M., Couve E., Lalloui L., Poyart C., Trieu-Cuot P.,  
 RA Kunst F.;  
 RT "Genome sequence of Streptococcus agalactiae, a pathogen causing  
 invasive neonatal disease.";  
 RL Mol. Microbiol. 45:1499-1513(2002).  
 DR EMBL; AL766850; CAD47009.1; --  
 DR SagalList; gbs1350; --

KW Hypothetical protein; Complete proteome.  
 SQ SEQUENCE 99 AA; 11302 MW; B83CB8FC1AB03C53 CRC64;  
 Query Match 40.8%; Score 31; DB 16; Length 99;  
 Best Local Similarity 46.2%; Pred. No. 8.4e+02;  
 Matches 6; Conservative 4; Mismatches 3; Indels 0; Gaps 0;  
 Qy 3 YDNLNVLMA 15  
 Db 78 YDNLNVLGIL 90  
 RESULT 29  
 Q8P2F3 PRELIMINARY; PRT; 100 AA.  
 ID Q8P2F3  
 AC Q8P2F3;  
 DT 01-OCT-2002 (TRENBLrel. 22, Created)  
 DT 01-OCT-2002 (TRENBLrel. 22, Last sequence update)  
 DT 01-OCT-2002 (TRENBLrel. 22, Last annotation update)  
 DE Hypothetical phage protein spym18\_0380 (Hypothetical phage associated  
 protein spym3\_0720).  
 GN SPYM18\_0380 OR SPYM3\_0720.  
 OS Streptococcus pyogenes (serotype M18), and  
 OS Streptococcus pyogenes (serotype M3).  
 OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;  
 OX NCBI\_TaxID=186103, 198466;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC SPYCIRS=S.pyogenes; STRAIN=MGAS8232 / Serotype M18;  
 RX MEDLINE=21927593; PubMed=11917108;  
 RA Smoot J.C., Barbican K.D., Van Gompel J.J., Smoot L.M., Chaussee M.S.,  
 RA Sylva G.L., Sturdevant D.E., Rickles S.M., Porcella S.F.,  
 RA Parkins L.D., Beres S.B., Campbell D.S., Smith T.M., Zhang Q.,  
 RA Kapur V., Daly J.A., Veasy L.G., Musser J.M.;  
 RT "Genome sequence and comparative microarray analysis of serotype M18  
 group A Streptococcus strains associated with acute rheumatic fever  
 outbreaks.";  
 RL Proc. Natl. Acad. Sci. U.S.A. 99:4668-4673(2002).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC SPYCIRS=S.pyogenes; STRAIN=MGAS315 / Serotype M3;  
 RX MEDLINE=22133808; PubMed=12122206;  
 RA Beres S.B., Sylva G.L., Barbican K.D., Lei B., Hoff J.S.,  
 RA Mammarella N.D., Liu M.-Y., Smoot J.C., Porcella S.F., Parkins L.D.,  
 RA Campbell D.S., Smith T.M., McCormick J.K., Leung D.Y.M.,  
 RA Schlievert P.M., Musser J.M.;  
 RT "Genome sequence of a serotype M3 strain of group A Streptococcus:  
 phage-encoded toxins, the high-virulence phenotype, and clone  
 emergence.";  
 RL Proc. Natl. Acad. Sci. U.S.A. 99:10078-10083(2002).  
 DR EMBL; AE009981; AAL97130.1; --  
 DR EMBL; AE014151; AAM79327.1; --  
 KW Hypothetical protein; Complete proteome.  
 SQ SEQUENCE 100 AA; 11394 MW; DBDB13AD8019BE6F CRC64;  
 Query Match 40.8%; Score 31; DB 16; Length 100;  
 Best Local Similarity 40.0%; Pred. No. 8.5e+02;  
 Matches 6; Conservative 4; Mismatches 5; Indels 0; Gaps 0;  
 Qy 2 YDNLNVLMA 16  
 Db 68 MYDKLNVYLNILA 82  
 RESULT 30  
 Q8HA80 PRELIMINARY; PRT; 89 AA.  
 ID Q8HA80  
 AC Q8HA80;  
 DT 01-MAR-2003 (TRENBLrel. 23, Created)  
 DT 01-MAR-2003 (TRENBLrel. 23, Last sequence update)  
 DT 01-MAR-2003 (TRENBLrel. 23, Last annotation update)  
 DE Hypothetical protein.

GN	ORF2C.	OC	Bacteria; Firmicutes; Bacillales; Staphylococcus.
OS	Bacteriophage T5.	OX	NCBI_TaxID=196620;
OC	Viruses; dsDNA viruses, no RNA stage; Caudovirales; Siphoviridae;	RN	[1]
OC	T5-like viruses.	RP	SEQUENCE FROM N.A.
OX	NCBI_TaxID=10726;	RX	MEDLINE=22040717; PubMed=12044378;
RN	[1]	RA	Baba T., Takeuchi F., Kuroda M., Yuzawa H., Aoki K.-I., Oguchi A.,
RP	SEQUENCE FROM N.A.	RA	Nagai Y., Iwama N., Asano K., Naimi T., Kuroda H., Cui L.,
RA	Mikoulinekaia G.V., Zimin A.A., Feofanov S.A.;	RA	Yamamoto K., Hiramatsu K.;
RL	Submitted (AUG-2002) to the EMBL/GenBank/DBJ databases.	RT	"Genome and virulence determinants of high virulence community-
RL	EMBL; AY140897; AANI7754.1; -	RT	acquired MRSA"; 1827(2002).
DR	Hypothetical protein.	RL	Lancet 359:1819-1827(2002).
KW	Hypothetical protein.	DR	EMBL; AP004826; BAB95046.1; -
SQ	SEQUENCE 89 AA; 10284 MW; 59C7EFOAC84D9B3C CRC64;	KW	Hypothetical protein; Complete proteome.
		SQ	SEQUENCE 29 AA; 3563 MW; 3B9E043A4272B619 CRC64;
Qy	1 RYDNLNV-LMAMNI 14	Query Match	39.5%; Score 30; DB 16; Length 29;
	::: ::	Best Local Similarity	55.6%; Pred. No. 3.8e+02;
Db	42 KKYDSLKNQVAFNL 56	Matches	7; Conservative 4; Mismatches 3; Indels 1; Gaps 1;
RESULT 31			
Q99UH5	PRELIMINARY; PRT; 29 AA.		
ID	Q99UH5	ID	Q9PPG6 PRELIMINARY; PRT; 35 AA.
AC	Q99UH5	AC	Q9PPG6
DT	01-JUN-2001 (TrEMBLrel. 17, Created)	DT	01-OCT-2000 (TrEMBLrel. 15, Created)
DT	01-JUN-2001 (TrEMBLrel. 17, Last sequence update)	DT	01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DE	01-MAR-2002 (TrEMBLrel. 20, Last annotation update)	DT	01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE	Hypothetical protein SAV1299.	DE	Hypothetical protein Cj0747.
GN	SAV1299 OR SALL139.1 OR SAS038.	GN	CJ0747.
OS	Staphylococcus aureus (strain N315).	OS	Campylobacter jejuni.
OS	Staphylococcus aureus (strain N315).	OC	Bacteria; Proteobacteria; Epsilonproteobacteria; Campylobacteriales;
OC	Bacteria; Firmicutes; Bacillales; Staphylococcus.	OC	Campylobacteraceae; Campylobacter.
OX	NCBI_TaxID=158878, 158879;	OX	NCBI_TaxID=197;
RN	[1]	RN	[1]
RP	SEQUENCE FROM N.A.	RP	SEQUENCE FROM N.A.
RC	SPECIES=S.aureus (strain Mu50), and S.aureus (strain N315);	RC	STRAIN=NCTC 11168;
RX	MEDLINE=21311952; PubMed=11418146;	RX	MEDLINE=20150912; PubMed=10688204;
RA	Kuroda M., Ohta T., Uchiyama I., Baba T., Yuzawa H., Kobayashi I.,	RA	Parkhill J., Wren B.W., Mungall K., Ketley J.M., Churcher C.,
RA	Cui L., Oguchi A., Aoki K.-I., Nagai Y., Lian J.-Q., Ito T.,	RA	Basham D., Chillingworth T., Davies R.M., Feltwell T., Holtroyd S.,
RA	Kanamori M., Matsumaru H., Maruyama A., Murakami H., Hosoyama A.,	RA	Jagels K., Karlyshev A.V., Moule S., Pallen M.J., Penn C.W.,
RA	Mizutani-Ui Y., Takahashi N.K., Sawano T., Inoue R.-I., Kaito C.,	RA	Quail M.A., Rajandream M.A., Rutherford K.M., van Vliet A.H.M.,
RA	Sekimizu K., Hirakawa H., Kuhara S., Goto S., Yabuzaki J.,	RA	Whitehead S., Barrrell B.G.;
RA	Kanehisa M., Yamashita A., Oshima K., Furuya K., Yoshino C., Shiba T.,	RT	"The genome sequence of the food-borne pathogen Campylobacter jejuni
RA	Hattori M., Ogasawara N., Hayashi H., Hiramatsu K.;	RT	reveals hypervariable sequences.";
RT	"Whole genome sequencing of methicillin-resistant Staphylococcus	RL	Nature 403:665-668(2000).
RL	aureus";	RL	EMBL; AL139076; CAB73017.1; -
DR	Lancet 357:1225-1240(2001).	DR	EMBL; AL139076; CAB73017.1; -
DR	EMBL; AP003361; BAB57461.1; -	KW	Hypothetical protein; Complete proteome.
DR	EMBL; AP003133; BAB42393.1; -	SQ	SEQUENCE 35 AA; 3912 MW; A1E0C2B43BB82277 CRC64;
KW	Hypothetical protein; Complete proteome.		
SQ	SEQUENCE 29 AA; 3562 MW; 3B9E043A462B619 CRC64;	Query Match	39.5%; Score 30; DB 16; Length 35;
		Best Local Similarity	25.9%; Pred. No. 4.6e+02;
		Matches	7; Conservative 5; Mismatches 1; Indels 14; Gaps 1;
Qy	2 VYDALNV-LMAMNI 10		
	::: ::		
Db	7 LYDTLNTLL 15		
RESULT 32			
Q8NW8	PRELIMINARY; PRT; 29 AA.		
ID	Q8NW8	ID	Q8ZV98 PRELIMINARY; PRT; 58 AA.
AC	Q8NW8	AC	Q8ZV98
DT	01-OCT-2002 (TrEMBLrel. 22, Created)	DT	01-MAR-2002 (TrEMBLrel. 20, Created)
DT	01-OCT-2002 (TrEMBLrel. 22, Last sequence update)	DT	01-MAR-2002 (TrEMBLrel. 20, Last sequence update)
DE	01-OCT-2002 (TrEMBLrel. 22, Last annotation update)	DE	Hypothetical protein PAE2389.
DE	Hypothetical protein MW1181.	OS	Pyrobaculum aerophilum.
GN	MW1181.	OS	Archaea; Crenarchaeota; Thermoprotei; Thermoproteales;
OS	Staphylococcus aureus (strain MW2).		

```

OC Thermoproteaceae; Pyrobaculum.
OX NCBI_TaxID=13773;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=IM2 / ATCC 51768 / DSM 7523;
RX PubMed=11792869;
RA Fitz-Gibbon S.T., Ladner H., Kim U.-J., Stetter K.O., Simon M.I.,
RA Miller J.H.;
RT "Genome sequence of the hyperthermophilic crenarchaeon Pyrobaculum
RT aerophilum.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:984-989(2002).
DR EMBL; AE009873; AA164158.1;
KW Hypothetical protein; Complete proteome.
SQ SEQUENCE 58 AA; 6342 MW; 110191142AD83792 CRC64;

Query Match 39.5%; Score 30; DB 17; Length 58;
Best Local Similarity 40.0%; Pred. No. 7.5e+02;
Matches 6; Conservative 3; Mismatches 6; Indels 0; Gaps 0;

Qy 1 RYDALNVLMMNII 15
Db 5 RKEGLNPFVAAGLI 19

RESULT 35
Q9KDN1 PRELIMINARY; PRT; 59 AA.
AC Q9KDN1;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 20, Last annotation update)
DE Hypothetical protein BH1180.
GN BH1180.
OS Bacillus halodurans.
OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.
OX NCBI_TaxID=86665;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C-125 / JCM 9153;
RX MEDLINE=20512582; PubMed=11058132;
RA Takami H., Nakasone K., Takaki Y., Maeno G., Sasaki R., Masui N.,
RA Fuji F., Hirano C., Nakamura Y., Ogawara N., Kuhara S.,
RA Horikoshi K.;
RT "Complete genome sequence of the alkaliphilic bacterium Bacillus
RT halodurans and genomic sequence comparison with Bacillus subtilis.";
RL Nucleic Acids Res. 28:4317-4331(2000).
DR EMBL; AP001511; BAB04899.1;
KW Hypothetical protein; Complete proteome.
SQ SEQUENCE 59 AA; 6857 MW; 8FE49D9F98915E77 CRC64;

Query Match 39.5%; Score 30; DB 16; Length 59;
Best Local Similarity 31.2%; Pred. No. 7.7e+02;
Matches 5; Conservative 6; Mismatches 5; Indels 0; Gaps 0;

Qy 1 RYDALNVLMMNII 16
Db 36 KVFRTINILMTILTS 51

RESULT 36
Q8F678 PRELIMINARY; PRT; 59 AA.
AC Q8F678;
DT 01-MAR-2003 (TrEMBLrel. 23, Created)
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE Hypothetical protein.
GN LA1431.
OS Leptospira interrogans.
OC Bacteria; Spirochaetes; Spirochaetales; Leptospiraceae; Leptospira.
OX NCBI_TaxID=173;
RN [1]
RP SEQUENCE FROM N.A.

```

```

RC STRAIN=56601 / Serogroup Icterohaemorrhagiae / Serovar lai;
RA Ren S.;
RL Submitted (MAR-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AE011323; AAN48630.1;
KW Hypothetical protein; Complete proteome.
SQ SEQUENCE 59 AA; 7098 MW; 5EC5757B6E80868D CRC64;

Query Match 39.5%; Score 30; DB 16; Length 59;
Best Local Similarity 41.7%; Pred. No. 7.7e+02;
Matches 5; Conservative 4; Mismatches 3; Indels 0; Gaps 0;

Qy 3 YDALNVLMMNII 14
Db 5 YDSINVSHSRNL 16

RESULT 37
Q9HRS3 PRELIMINARY; PRT; 60 AA.
AC Q9HRS3;
DT 01-MAR-2001 (TrEMBLrel. 16, Created)
DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT 01-MAR-2001 (TrEMBLrel. 16, Last annotation update)
DE VNG0568C.
GN VNG0568C.
OS Halobacterium sp. (strain NRC-1).
OC Archaea; Euryarchaeota; Halobacteria; Halobacteriales;
OC Halobacteriaceae; Halobacterium.
OX NCBI_TaxID=64091;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=20504483; PubMed=11016950;
RA Ng W.V., Kennedy S.P., Mahairas G.G., Berquist B., Pan M.,
RA Shukla H.D., Lasky S.R., Balliga N.S., Thorsson V., Sbrogna J.,
RA Swartzell S., Weir D., Hall J., Dahl T.A., Welti R., Goo Y.A.,
RA Leithausen B., Keller K., Cruz R., Banson M.J., Hough D.W., Dale H.,
RA Maddocks D.G., Jablonski P.E., Krebs M.P., Angevine C.M., Jung K.-H.,
RA Isenbarger T.A., Peck R.F., Pohlshocher M., Spudich J.L., Omer A.D.,
RA Alam M., Freitas T., Hou S., Daniels C.J., Dennis P.P., Omer A.D.,
RA Ebhardt H., Lowe T.M., Liang P., Riley M., Hood L., DasSarma S.;
RT "Genome sequence of Halobacterium species NRC-1.";
RL Proc. Natl. Acad. Sci. U.S.A. 97:12176-12181(2000).
DR EMBL; AE005007; AAG19085.1;
KW Complete proteome.
SQ SEQUENCE 60 AA; 6244 MW; 55D3A62675041011 CRC64;

Query Match 39.5%; Score 30; DB 17; Length 60;
Best Local Similarity 53.8%; Pred. No. 7.8e+02;
Matches 7; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

Qy 2 VYDALNVLMMNII 14
Db 37 VYALNFMISIAI 49

RESULT 38
Q8GFK0 PRELIMINARY; PRT; 64 AA.
AC Q8GFK0;
DT 01-MAR-2003 (TrEMBLrel. 23, Created)
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE ORF39.
OS Staphylococcus aureus.
OC Plasmid EDINA plasmid.
OC Bacteria; Firmicutes; Bacillales; Staphylococcus.
OX NCBI_TaxID=1280;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=E-1;
RA Sugai M., Yamaguchi T., Hayashi T., Nakasone K., Takami H.;
RT "Complete nucleotide sequence of Staphylococcus aureus E-1 EDINA
RT plasmid.";

```

Tue Feb 17 11:55:53 2004

RL Submitted (JAN-2001) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; AP003089; BAC54531.1; -  
 KW Plasmid.  
 SQ SEQUENCE 64 AA; 7316 MW; 59BP2A1EEAE9A7F9 CRC64;  
  
 Query Match 39.5%; Score 30; DB 2; Length 64;  
 Best Local Similarity 45.5%; Pred. No. 8.3e+02;  
 Matches 5; Conservative 5; Mismatches 1; Indels 0; Gaps 0;  
  
 Qy 3 YDALNVLMMANN 13  
 Db 41 YPSLDILIAAMS 51  
  
 RESULT 39  
 Q91FU2 PRELIMINARY; PRT; 64 AA.  
 AC Q91FU2;  
 DT 01-DEC-2001 (TrEMBLrel. 19, Created)  
 DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)  
 DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)  
 DE 230L.  
 OS Chilo iridescent virus (CIV) (Insect iridescent virus type 6).  
 OC Viruses; dsDNA viruses, no RNA stage; Iridoviridae; Iridovirus.  
 OX NCBI\_TaxID=10488;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Delius H., Darai G., Fluegel R.M.;  
 RT "DNA analysis of insect iridescent virus 6: evidence for circular  
 permatation and terminal redundancy.";  
 RL J. Virol. 49:609-614 (1984).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=86174607; PubMed=3959991;  
 RA Lorbacher de Ruiz H., Gelderblom H., Hofmann W., Darai G.;  
 RT "Insect iridescent virus type 6 induced toxic degenerative hepatitis  
 in mice.";  
 RL Med. Microbiol. Immunol. 175:43-53 (1986).  
 RN [3]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=87321126; PubMed=2820141;  
 RA Schnitzler P., Soltan J.B., Fischer M., Reinsner H., Scholz J.,  
 RA Delius H., Darai G.;  
 RT "Molecular cloning and physical mapping of the genome of insect  
 iridescent virus type 6: further evidence for circular permutation of  
 the viral genome.";  
 RL Virology 160:66-74 (1987).  
 RN [4]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=89073752; PubMed=3201750;  
 RA Fischer M., Schnitzler P., Delius H., Darai G.;  
 RT "Identification and characterization of the repetitive DNA element in  
 the genome of insect iridescent virus type 6.";  
 RL Virology 167:485-496 (1988).  
 RN [5]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=92196996; PubMed=1549908;  
 RA Handermann M., Schnitzler P., Rosen-Wolff A., Raab K., Sonntag K.C.,  
 RA Darai G.;  
 RT "Identification and mapping of origins of DNA replication within the  
 DNA sequences of the genome of insect iridescent virus type 6.";  
 RL Virus Genes 6:19-32 (1992).  
 RN [6]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=93118242; PubMed=1475907;  
 RA Sonntag K.C., Darai G.;  
 RT "Characterization of the third origin of DNA replication of the genome  
 of insect iridescent virus type 6.";  
 RL Virus Genes 6:333-342 (1992).  
 RN [7]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=93260401; PubMed=8492091;  
 RA Stohwasser R., Raab K., Schnitzler P., Janssen W., Darai G.;  
 RT "Identification of the gene encoding the major capsid protein of  
 insect iridescent virus type 6 by polymerase chain reaction.";  
 J. Gen. Virol. 74:873-879 (1993).  
 RN [8]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=94167241; PubMed=8121799;  
 RA Schnitzler P., Hug M., Handermann M., Janssen W., Koonin E.V.,  
 RA Delius H., Darai G.;  
 RT "Identification of genes encoding zinc finger proteins, non-histone  
 chromosomal HMG protein homologue, and a putative GTP phosphohydrolase  
 in the genome of Chilo iridescent virus.";  
 Nucleic Acids Res. 22:158-166 (1994).  
 RN [9]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=94353641; PubMed=8073636;  
 RA Sonntag K.C., Schnitzler P., Koonin E.V., Darai G.;  
 RT "Chilo iridescent virus encodes a putative helicase belonging to a  
 distinct family within the 'DEAD/H' superfamily: implications for the  
 evolution of large DNA viruses.";  
 RL Virus Genes 8:151-158 (1994).  
 RN [10]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=95213160; PubMed=7698884;  
 RA Sonntag K.C., Schnitzler P., Janssen W., Darai G.;  
 RT "Identification of the primary structure and the coding capacity of  
 the genome of insect iridescent virus type 6 between the genome  
 coordinates 0.310 and 0.347 (7990 bp).";  
 RL Intervirology 37:287-297 (1994).  
 RN [11]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=94292906; PubMed=8021587;  
 RA Schnitzler P., Sonntag K.C., Muller M., Janssen W., Bugert J.J.,  
 RA Koonin E.V., Darai G.;  
 RT "Insect iridescent virus type 6 encodes a polypeptide related to the  
 largest subunit of eukaryotic RNA polymerase II.";  
 RL J. Gen. Virol. 75:1557-1567 (1994).  
 RN [12]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=98141693; PubMed=9482589;  
 RA Bahr U., Tidona C.A., Darai G.;  
 RT "The DNA sequence of Chilo iridescent virus between the genome  
 coordinates 0.101 and 0.391: similarities in coding strategy between  
 insect and vertebrate iridoviruses.";  
 RL Virus Genes 15:235-245 (1997).  
 RN [13]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=99125223; PubMed=9926400;  
 RA Muller K., Tidona C.A., Bahr U., Darai G.;  
 RT "Identification of a thymidylate synthase gene within the genome of  
 Chilo iridescent virus.";  
 RL Virus Genes 17:243-258 (1998).  
 RN [14]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=99383793; PubMed=10456793;  
 RA Muller K., Tidona C.A., Darai G.;  
 RT "Identification of a gene cluster within the genome of Chilo  
 iridescent virus encoding enzymes involved in viral DNA replication  
 and processing.";  
 RL Virus Genes 18:243-264 (1999).  
 RN [15]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=21342589; PubMed=11448171;  
 RA Jakob N.J., Muller K., Bahr U., Darai G.;  
 RT "Analysis of the First Complete DNA Sequence of an Invertebrate  
 Iridovirus: Coding Strategy of the Genome of Chilo Iridescens Virus.";  
 RL Virology 286:182-196 (2001).  
 RN [16]  
 RP SEQUENCE FROM N.A.  
 RA Jakob N.J., Muller K., Bahr U., Darai G.;  
 RL Submitted (SEP-2000) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; AF303741; AAK82091.1; -  
 SQ SEQUENCE 64 AA; 7741 MW; 42CB68481A85F957 CRC64;

```

Query Match      39.5%; Score 30; DB 12; Length 64;
Best Local Similarity 45.5%; Pred. No. 8.3e+02;
Matches 5; Conservative 5; Mismatches 1; Indels 0; Gaps 0;

QY      5 ALNVLMMNII 15
      |||:|:|:|
Db      31 AIDLVLNII 41

RESULT 40
Q9GMS9          PRELIMINARY;      PRT;      71 AA.
AC      Q9GMS9;
DT      01-MAR-2001 (TrEMBLrel. 16, Created)
DT      01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT      01-MAR-2001 (TrEMBLrel. 16, Last annotation update)
DE      Hypothetical 8.3 kDa protein.
OS      Macaca fascicularis (Crab eating macaque) (Cynomolgus monkey).
OC      Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC      Mammalia; Eutheria; Primates; Catarrhini; Cercopitheidae;
OC      Cercopitheinae; Macaca.
OX      NCBI_TaxID=9541;
RN      [1]
RP      SEQUENCE FROM N.A.
RC      TISSUE=Cerebellum;
RA      Osada N., Hida M., Kusuda J., Tanuma R., Iseki K., Hirai M., Terao K.,
RA      Suzuki Y., Sugano S., Hashimoto K.;
RT      "Isolation of full-length cDNA clones from macaque brain cDNA
RT      libraries.";
RL      Submitted (AUG-2000) to the EMBL/GenBank/DBJ databases.
DR      EMBL; AB047839; BAB12265.1; -.
KW      Hypothetical protein.
SQ      SEQUENCE 71 AA; 8296 MW; 1A50C4AE703A69C7 CRC64;

Query Match      39.5%; Score 30; DB 6; Length 71;
Best Local Similarity 35.7%; Pred. No. 9.2e+02;
Matches 5; Conservative 4; Mismatches 5; Indels 0; Gaps 0;

QY      1 RYDALNVLMMNI 14
      :|:|:|:|:|
Db      10 KLYDNINVKLSQV 23

RESULT 41
Q9PDS4          PRELIMINARY;      PRT;      72 AA.
AC      Q9PDS4;
DT      01-OCT-2000 (TrEMBLrel. 15, Created)
DT      01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT      01-MAR-2002 (TrEMBLrel. 20, Last annotation update)
DE      Hypothetical protein Xf1305.
GN      Xf1305.
OS      Xylella fastidiosa.
OC      Bacteria; Proteobacteria; Gammaproteobacteria; Xanthomonadales;
OC      Xanthomonadaceae; Xylella.
OX      NCBI_TaxID=2371;
RN      [1]
RP      SEQUENCE FROM N.A.
RC      STRAIN=9a5C;
RX      MEDLINE=20365717; PubMed=10910347;
RA      Simpson A.J.G., Reinach F.C., Arruda P., Abreu F.A., Acencio M.,
RA      Alvarenga R., Alves L.M.C., Araya J.E., Baia G.S., Baptista C.S.,
RA      Barros M.H., Bonaccorsi E.D., Bordin S., Bove J.M., Briones M.R.S.,
RA      Bueno M.R.P., Camargo A.A., Camargo L.E.A., Carraro D.M., Carrier H.,
RA      Coutinho L.B., Colombo C., Costa F.F., Costa M.C.R., Costa-Neto C.M.,
RA      Facincani A.P., Ferreira A.J.S., Ferreira V.C.A., Ferro J.A.,
RA      Fraga J.S., Franca S.C., Franco M.C., Frohme M., Furlan L.R.,
RA      Garnier M., Goldman G.H., Goldman M.H.S., Gomes S.L., Gruber A.,
RA      Ho P.L., Hoheisel J.D., Junqueira M.L., Kemper E.L., Kitajima J.P.,
RA      Krieger J.E., Kuramae E.E., Laigret F., Lambais M.R., Leite L.C.C.,
RA      Lemos E.G.M., Lemos M.V.F., Lopes S.A., Lopes C.R., Machado J.A.,
RA      Machado M.A., Madeira A.M.B.N., Madeira H.M.F., Marino C.L.,

```

```

RA      Marques M.V., Martins E.A.L., Martins E.M.F., Matsukuma A.Y.,
RA      Menck C.F.M., Miracca E.C., Miyaki C.Y., Monteiro-Vitorello C.B.,
RA      Moon D.H., Nagai M.A., Nascimento A.L.T.O., Netto L.E.S.,
RA      Nhani A., Jr., Nobrega F.G., Nunes L.R., Oliveira M.A.,
RA      de Oliveira M.C., de Oliveira R.C., Palmieri D.A., Paris A.,
RA      Peixoto B.R., Pereira G.A.G., Pereira H.A. Jr., Pesquero J.B.,
RA      Quaggio R.B., Roberto P.G., Rodrigues V., de Rosa A.J.M.,
RA      de Rosa V.E. Jr., de Sa R.G., Santelli R.V., Sawasaki H.E.,
RA      da Silva A.C.R., da Silva A.M., da Silva F.R., Silva W.A. Jr.,
RA      da Silveira J.F., Silvestri M.L.Z., Siqueira W.J., de Souza A.A.,
RA      de Souza A.P., Terenzi M.F., Truffi D., Tsai S.M., Tsuchioka M.H.,
RA      Vallada H., Van Sluys M.A., Verjovski-Almeida S., Vettore A.L.,
RA      Zago M.A., Zatz M., Meidanis J., Setubal J.C.;
RT      "The genome sequence of the plant pathogen Xylella fastidiosa.";
RL      Nature 406:151-159(2000).
EMBL; AE003963; AAF84114.1; -.
KW      Hypothetical protein; Complete proteome.
SQ      SEQUENCE 72 AA; 7462 MW; 1855A508694610B5 CRC64;

Query Match      39.5%; Score 30; DB 16; Length 72;
Best Local Similarity 66.7%; Pred. No. 9.3e+02;
Matches 6; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY      2 YDNLNVLMM 10
      |||:|:|
Db      53 VYDAFNSFM 61

RESULT 42
Q9AC63          PRELIMINARY;      PRT;      79 AA.
AC      Q9AC63;
DT      01-JUN-2001 (TrEMBLrel. 17, Created)
DT      01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT      01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE      Hypothetical protein SAP028.
GN      SAP028.
OS      Staphylococcus aureus (strain N315).
OG      Plasmid pN315.
OC      Bacteria; Firmicutes; Bacillales; Staphylococcus.
OX      NCBI_TaxID=158879;
RN      [1]
RP      SEQUENCE FROM N.A.
RX      MEDLINE=21311952; PubMed=11418146;
RA      Kuroda M., Ohta T., Uchiyama I., Baba T., Yuzawa H., Kobayashi I.,
RA      Cui L., Oguchi A., Aoki K.I., Nagai Y., Ito T., Kanamori M.,
RA      Matsuura H., Maruyama A., Murakami H., Hosoyama A., Mizutani-Ui Y.,
RA      Takahashi N.K., Sawano T., Inoue R.I., Kaito C., Sekimizu K.,
RA      Hirakawa H., Kuhara S., Goto S., Yabuzaki J., Kanehisa M.,
RA      Yamashita A., Oshima K., Furuya K., Yoshino C., Shiba T., Hattori M.,
RA      Ogasawara N., Hayashi H., Hiramatsu K.;
RT      "Whole genome sequencing of methicillin-resistant Staphylococcus
RT      aureus.";
RL      Lancet 357:1225-1240(2001).
DR      EMBL; AP003139; BAB43897.1; -.
DR      InterPro; IPR001387; HTH_3.
DR      Pfam; PF01381; HTH_3; 1.
DR      SMART; SM00530; HTH_XRE; 1.
KW      Plasmid; Hypothetical protein; Complete proteome.
SQ      SEQUENCE 79 AA; 9192 MW; D019DADDAACE5D52 CRC64;

Query Match      39.5%; Score 30; DB 16; Length 79;
Best Local Similarity 45.5%; Pred. No. 1e+03;
Matches 5; Conservative 5; Mismatches 1; Indels 0; Gaps 0;

QY      3 YDNLNVLMMN 13
      |:|:|:|:|
Db      41 YPSLDLIIAMS 51

RESULT 43
Q98215          PRELIMINARY;      PRT;      82 AA.
ID      Q98215

```

```

AC Q98215;
DT 01-FEB-1997 (TrEMBLrel. 02, Created)
DT 01-FEB-1997 (TrEMBLrel. 02, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE MCO47L.
GN Molluscum contagiosum virus subtype 1 (MCV1).
OS Viruses; deDNA viruses, no RNA stage; Poxviridae; Chordopoxvirinae;
OC Molluscipoxvirus.
OC NCBI_TaxID=10280;
OX [1]
RN SEQUENCE FROM N.A.
RP MEDLINE=96325459; PubMed=8670425;
RX Senkevich T.G., Bugert J.J., Sisler J.R., Koonin E.V., Darai G.,
PA Moss B., sequence of a human tumorigenic poxvirus: Prediction of
RT specific host response-evasion genes.";
RL Science 273:813-816(1996).
RN [2]
RP SEQUENCE FROM N.A.
RX Senkevich T.G., Bugert J.J., Sisler J.R., Koonin E.V., Darai G.,
PA Moss B.;
RL Submitted (JUN-1996) to the EMBL/GenBank/DBJ databases.
DR EMBL; U60315; AAC55175.1; -.
DR InterPro; IPR006803; Pox 15.
DR Pfam; PF04713; Pox 15; 1.
SQ SEQUENCE 82 AA; 8606 MW; 740DE5C73803B4AB CRC64;

Query Match 39.5%; Score 30; DB 12; Length 82;
Best Local Similarity 50.0%; Pred. No. 1.1e+03;
Matches 6; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

QY 5 ALNVLMMNIIIS 16
DB 13 ALTVMMLMWS 24

RESULT 44
Q8RUQ9 PRELIMINARY; PRT; 92 AA.
AC Q8RUQ9;
DT 01-JUN-2002 (TrEMBLrel. 21, Created)
DT 01-JUN-2002 (TrEMBLrel. 21, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE Acetyl-CoA carboxylase (Fragment).
OS Lolium rigidum.
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Pooideae;
OC Poaceae; Lolium.
OC NCBI_TaxID=89674;
OX [1]
RN SEQUENCE FROM N.A.
RP Tal A., Rubin B.;
RT "A point mutation in the ACCase gene is responsible for the resistance
to ACCase inhibiting herbicides in Lolium rigidum.";
RL Submitted (FEB-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF482471; AAL87701.1; -.
DR EMBL; AF482473; AAL87703.1; -.
DR InterPro; IPR000022; Carboxyl_trans.
DR Pfam; PF01039; Carboxyl_trans; 1.
DR NON_TER 1
FT NON_TER 92
SQ SEQUENCE 92 AA; 10033 MW; 92ABFA9E0B2BF946 CRC64;

Query Match 39.5%; Score 30; DB 10; Length 92;
Best Local Similarity 42.9%; Pred. No. 1.2e+03;
Matches 6; Conservative 2; Mismatches 6; Indels 0; Gaps 0;

QY 3 YDNLVLMNMNIIIS 16
DB 76 YSALNKLGLREVYS 89

RESULT 45
Q8S3M1 PRELIMINARY; PRT; 92 AA.
AC Q8S3M1;
DT 01-JUN-2002 (TrEMBLrel. 21, Created)
DT 01-JUN-2002 (TrEMBLrel. 21, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE Acetyl-CoA carboxylase (Fragment).
OS Lolium rigidum.
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Pooideae;
OC Poaceae; Lolium.
OC NCBI_TaxID=89674;
OX [1]
RN SEQUENCE FROM N.A.
RP Tal A., Rubin B.;
RT "A point mutation in the ACCase gene is responsible for the resistance
to ACCase inhibiting herbicides in Lolium rigidum.";
RL Submitted (FEB-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF482472; AAL87702.1; -.
DR InterPro; IPR000022; Carboxyl_trans.
DR Pfam; PF01039; Carboxyl_trans; 1.
DR NON_TER 1
FT NON_TER 92
SQ SEQUENCE 92 AA; 10033 MW; D3AFACF0B2BFE31 CRC64;

Query Match 39.5%; Score 30; DB 10; Length 92;
Best Local Similarity 42.9%; Pred. No. 1.2e+03;
Matches 6; Conservative 2; Mismatches 6; Indels 0; Gaps 0;

QY 3 YDNLVLMNMNIIIS 16
DB 76 YSALNKLGLREVYS 89

Search completed: February 17, 2004, 10:56:22
Job time : 23.6535 secs

```

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: February 17, 2004, 10:50:13 ; Search time 5.06931 Seconds  
(without alignments)  
148.428 Million cell updates/sec

Title: US-09-900-147-5

Perfect score: 76  
Sequence: 1 RYVDALNVLMMNNIIS 16

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 127863 seqs, 47026705 residues

Total number of hits satisfying chosen parameters: 13973

Minimum DB seq length: 0

Maximum DB seq length: 100

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt\_41.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	33	43.4	38	1 E2F1_RAT	O09139 rattus norv
2	30	39.5	56	1 SCF2_MESMA	Q91397 mesobuthus
3	30	39.5	56	1 SCF3_MESMA	Q948d1 mesobuthus
4	30	39.5	96	1 YFRC_PROVU	P20927 proteus vul
5	29	38.2	70	1 YORA_ITV1	P19285 thermoprote
6	29	38.2	98	1 NULM_DASNO	O21333 dasypus nov
7	28	36.8	42	1 RET5_BOVIN	P82708 bos taurus
8	28	36.8	79	1 V105_VACCC	P20500 vaccinia vi
9	28	36.8	79	1 V105_VACCV	P12924 vaccinia vi
10	28	36.8	79	1 V105_VACV	P33001 variola vir
11	28	36.8	83	1 PSBE_NEPOL	Q9tky1 nephroselmi
12	27	35.5	48	1 ATP8_CANPA	P17345 candida par
13	27	35.5	58	1 YC18_FORPU	P51366 porphyra pu
14	27	35.5	91	1 ILG1_CAEEL	Q18060 caenorhabdi
15	27	35.5	91	1 VAPD_HAEIN	P71351 haemophilus
16	27	35.5	93	1 FLJE_AQUAE	O67242 aquifex aeo
17	27	35.5	95	1 VAPD_ACTAC	Q52243 actinobacil
18	27	35.5	98	1 NULM_BALMU	P41301 balaeonopter
19	26.5	34.9	77	1 ACPH_UREPA	Q9ppv4 ureaplasma
20	26	34.2	48	1 YX93_CYPAP	P48330 cyanophora
21	26	34.2	56	1 RL32_ECOLI	P02435 escherichia
22	26	34.2	67	1 ATP1_STRMU	P95783 streptococc
23	26	34.2	67	1 YIFL_ECOLI	P39166 escherichia
24	26	34.2	69	1 Y010_TROWT	P59486 tropheryma
25	26	34.2	81	1 PSK6_ARATH	Q81a14 arabidopsis
26	26	34.2	89	1 FLIQ_ERWCA	P34201 erwinia car
27	26	34.2	89	1 TRP6_BOVIN	Q9myw0 bos taurus
28	26	34.2	89	1 VDMA_BPT3	P10309 bacterioph
29	26	34.2	89	1 VDMA_BPT7	P36393 bacterioph
30	26	34.2	90	1 EFIB_AERPE	Q99804 aeropyrum p
31	26	34.2	91	1 NULM_RHISA	O99826 rhipicephal
32	26	34.2	92	1 IAPP_CAVPO	P12966 cavia porce
33	26	34.2	92	1 YCHH_ECOLI	P31807 escherichia

34 26 34.2 96 1 C553\_HELPJ Q9zj29 helicobacte  
35 26 34.2 98 1 NULM\_BALPH P24976 balaeonopter  
36 25.5 33.6 98 1 RK23\_ASTLO P34771 astasia lon  
37 25 32.9 33 1 YC12\_CHLRE P50370 chlamydomon  
38 25 32.9 48 1 VGR\_BPNF Q04395 bacterioph  
39 25 32.9 52 1 LHA1\_ECTHL P80102 ectothiorho  
40 25 32.9 56 1 Y546\_METUA Q57966 methanococc  
41 25 32.9 57 1 SCF1\_MESMA Q948d2 mesobuthus  
42 25 32.9 70 1 COXY\_YEAST P3103 saccharomyc  
43 25 32.9 70 1 YNCH\_ECOLI P76118 escherichia  
44 25 32.9 80 1 DBHL\_RICCN Q92h14 rickettsia  
45 25 32.9 82 1 RANA\_RANFI P08950 rana pipien

#### ALIGNMENTS

##### RESULT 1

E2F1\_RAT ID E2F1\_RAT STANDARD; PRT; 38 AA.  
AC O09139;  
DT 01-NOV-1997 (Rel. 35, Created)  
DT 01-NOV-1997 (Rel. 35, Last sequence update)  
DT 15-JUL-1998 (Rel. 36, Last annotation update)  
DE Transcription factor E2F1 (E2F-1) (Fragment).  
GN E2F1.  
OS Rattus norvegicus (Rat).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.  
OX NCBI\_TaxID=10116;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=96113578; PubMed=8673024;  
RA Hosokawa Y., Yang M., Kaneko S., Tanaka M., Nakashima K.;  
RT "Synergistic gene expressions of cyclin E, cdk2, cdk5 and E2F-1  
RT during the prolactin-induced G1/S transition in rat Nd2 pre-T  
RT lymphoma cells";  
RL Biochem. Mol. Biol. Int. 37:393-399(1995).  
CC -!- FUNCTION: TRANSCRIPTIONAL ACTIVATOR THAT BINDS TO E2F SITES. THESE  
CC SITES ARE PRESENT IN THE PROMOTER OF MANY GENES WHOSE PRODUCTS ARE  
CC INVOLVED IN CELL PROLIFERATION. MAY MEDIATE GROWTH FACTOR-  
CC INITIATED SIGNAL TRANSDUCTION (BY SIMILARITY).  
CC -!- SUBUNIT: COMPONENT OF THE DRTF1/E2F TRANSCRIPTION FACTOR COMPLEX.  
CC BINDS COOPERATIVELY WITH DP-1 TO E2F SITES. INTERACTS  
CC PREFERENTIALLY WITH RETINOBLASTOMA PROTEIN RB1 THAT INHIBIT THE  
CC E2F TRANSACTIVATION DOMAIN. SOME INTERACTION HAS ALSO BEEN FOUND  
CC WITH RETINOBLASTOMA RELATED PROTEIN 107 (BY SIMILARITY).  
CC -!- SUBCELLULAR LOCATION: Nuclear.  
CC -!- SIMILARITY: BELONGS TO THE E2F/DP FAMILY.

-----  
This SWISS-PROT entry is copyright. It is produced through a collaboration  
between the Swiss Institute of Bioinformatics and the EMBL outstation -  
the European Bioinformatics Institute. There are no restrictions on its  
use by non-profit institutions as long as its content is in no way  
modified and this statement is not removed. Usage by for commercial  
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
-----

EMBL; D63165; BAA09641.1; -;  
HSSP; Q16254; ICF7.  
DR TRANSFAC; T02952; -;  
DR InterPro; IPR003316; E2F TDP.  
DR Pfam; PF02319; E2F TDP; 1.  
KW Transcription regulation; Activator; DNA-binding; Nuclear protein.  
FT NON\_TER 1  
FT DOMAIN <1>38 DNA-BINDING (POTENTIAL).  
FT NON\_TER 38 38  
SQ SEQUENCE 38 AA; 4313 MW; E7903458375CC9E4 CRC64;

Query Match 43.4%; Score 33; DB 1; Length 38;  
Best Local Similarity 66.7%; Pred. No. 15;  
Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 RYVDALNVL 9  
DB 28 RIYDITNLV 36

RESULT 2  
SCP2\_MESMA STANDARD; PRT; 56 AA.

ID SCP2\_MESMA  
AC Q9NJ7P; PS8491;  
DT 28-FEB-2003 (Rel. 41, Last sequence update)  
DE Mesobuthus martensii (Manchurian scorpion) (Buthus martensii).  
OC Eukaryota; Metazoa; Arthropoda; Chelicerata; Arachnida; Scorpiones;  
OC Buthoidea; Buthidae; Mesobuthus.  
NCBI\_TaxID=34649;

[1]  
SEQUENCE FROM N.A.  
RP TISSUE=Venom gland;  
RC MEDLINE=99313193; PubMed=10386622;  
RA Wu J.-J., Dai L., Lan Z.D., Chi C.-W.;  
RT "Genomic organization of three neurotoxins active on small conductance  
DE Neurotoxin Bmp02 precursor (potassium ion channel blocker P02P).  
OS Mesobuthus martensii (Manchurian scorpion) (Buthus martensii).  
OC Eukaryota; Metazoa; Arthropoda; Chelicerata; Arachnida; Scorpiones;  
OC Buthoidea; Buthidae; Mesobuthus.  
NCBI\_TaxID=34649;

[2]  
SEQUENCE FROM N.A.  
RP TISSUE=Venom gland;  
RC MEDLINE=99402983; PubMed=10471839;  
RA Zhu S.-Y., Li W.-X., Zeng X.-C., Jiang D.-H., Mao X., Liu H.;  
RT "Molecular cloning and sequencing of two 'short chain' and two 'long  
chain' K(+) channel-blocking peptides from the Chinese scorpion Buthus  
DE Martensii Karsch.";  
OS Martensii Karsch.";  
OC Buthoidea; Buthidae; Mesobuthus.  
NCBI\_TaxID=34649;

[3]  
STRUCTURE BY NMR OF 29-56.  
RP TISSUE=Venom;  
RC MEDLINE=20530297; PubMed=11076505;  
RA Xu Y.Q., Wu J.H., Pei J.M., Shi Y.Y., Ji Y.H., Tong Q.C.;  
RT "Solution structure of Bmp02, a new potassium channel blocker from the  
DE Venom of the Chinese scorpion Buthus martensii Karsch.";  
OS Blochemistry 39:13669-13675(2000).  
CC -1- FUNCTION: Blocks potassium channels.  
CC -1- SUBCELLULAR LOCATION: Secreted.  
CC -1- TISSUE SPECIFICITY: Expressed by the venom gland.  
CC -1- SIMILARITY: BELONGS TO THE SCORPION LEIUTOXIN FAMILY.

This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL Outstation at the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/ or send an email to license@isb-sib.ch).

EMBL; AF132975; AAF31296.1; --  
PDB; 1DU9; 28-MAR-01.  
KW Toxin; Neurotoxin; Ionic channel inhibitor;  
KW Potassium channel inhibitor; Signal; 3D-structure.  
FT SIGNAL 1 28  
FT CHAIN 29 56 NEUROTOXIN BMP02.  
FT DISULFID 31 47  
FT DISULFID 34 52  
FT DISULFID 38 54  
SQ SEQUENCE 56 AA; 6015 MW; 70953032042F8672 CRC64;

Query Match 39.5%; Score 30; DB 1; Length 56;  
Best Local Similarity 26.7%; Pred. No. 80;  
Matches 4; Conservative 8; Mismatches 3; Indels 0; Gaps 0;

QY 1 RYVDALNVLMMNNII 15  
DB 3 RLFTLVLLVLMNMV 17

RESULT 3  
SCP3\_MESMA STANDARD; PRT; 56 AA.

ID SCP3\_MESMA  
AC Q9USDL; PS8491;  
DT 28-FEB-2003 (Rel. 41, Created)  
DE Mesobuthus martensii (Manchurian scorpion) (Buthus martensii).  
OC Eukaryota; Metazoa; Arthropoda; Chelicerata; Arachnida; Scorpiones;  
OC Buthoidea; Buthidae; Mesobuthus.  
NCBI\_TaxID=34649;

[1]  
SEQUENCE FROM N.A.  
RP TISSUE=Venom gland;  
RC MEDLINE=99402983; PubMed=10471839;  
RA Zhu S.-Y., Li W.-X., Zeng X.-C., Jiang D.-H., Mao X., Liu H.;  
RT "Molecular cloning and sequencing of two 'short chain' and two 'long  
chain' K(+) channel-blocking peptides from the Chinese scorpion Buthus  
DE Martensii Karsch.";  
OS Martensii Karsch.";  
OC Buthoidea; Buthidae; Mesobuthus.  
NCBI\_TaxID=34649;

[2]  
SUBCELLULAR LOCATION: Secreted.  
CC -1- TISSUE SPECIFICITY: Expressed by the venom gland.  
CC -1- SIMILARITY: BELONGS TO THE SCORPION LEIUTOXIN FAMILY.

This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL Outstation at the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/ or send an email to license@isb-sib.ch).

EMBL; AF132975; AAF31296.1; --  
PDB; 1DU9; 28-MAR-01.  
KW Toxin; Neurotoxin; Ionic channel inhibitor;  
KW Potassium channel inhibitor; Signal; 3D-structure.  
FT SIGNAL 1 28  
FT CHAIN 29 56 NEUROTOXIN BMP02.  
FT DISULFID 31 47  
FT DISULFID 34 52  
FT DISULFID 38 54  
SQ SEQUENCE 56 AA; 6015 MW; 70953032042F8672 CRC64;

Query Match 39.5%; Score 30; DB 1; Length 56;  
Best Local Similarity 26.7%; Pred. No. 80;  
Matches 4; Conservative 8; Mismatches 3; Indels 0; Gaps 0;

QY 1 RYVDALNVLMMNNII 15  
DB 3 RLFTLVLLVLMNMV 17

RESULT 4  
YFCR\_PROVU STANDARD; PRT; 96 AA.

ID YFCR\_PROVU  
AC P20927; 1991 (Rel. 17, Created)  
DT 01-FEB-1991 (Rel. 17, Last sequence update)  
DE Mesobuthus martensii (Manchurian scorpion) (Buthus martensii).  
OC Eukaryota; Metazoa; Arthropoda; Chelicerata; Arachnida; Scorpiones;  
OC Buthoidea; Buthidae; Mesobuthus.  
NCBI\_TaxID=34649;

[1]  
SEQUENCE FROM N.A.  
RP TISSUE=Venom gland;  
RC MEDLINE=99313193; PubMed=10386622;  
RA Wu J.-J., Dai L., Lan Z.D., Chi C.-W.;  
RT "Genomic organization of three neurotoxins active on small conductance  
DE Neurotoxin Bmp03 precursor (potassium ion channel blocker P03P).  
OS Mesobuthus martensii (Manchurian scorpion) (Buthus martensii).  
OC Eukaryota; Metazoa; Arthropoda; Chelicerata; Arachnida; Scorpiones;  
OC Buthoidea; Buthidae; Mesobuthus.  
NCBI\_TaxID=34649;

[2]  
SEQUENCE FROM N.A.  
RP TISSUE=Venom gland;  
RC MEDLINE=99402983; PubMed=10471839;  
RA Zhu S.-Y., Li W.-X., Zeng X.-C., Jiang D.-H., Mao X., Liu H.;  
RT "Molecular cloning and sequencing of two 'short chain' and two 'long  
chain' K(+) channel-blocking peptides from the Chinese scorpion Buthus  
DE Martensii Karsch.";  
OS Martensii Karsch.";  
OC Buthoidea; Buthidae; Mesobuthus.  
NCBI\_TaxID=34649;

[3]  
FUNCTION: Blocks small conductance Calcium-activated potassium channels.  
CC -1- SUBCELLULAR LOCATION: Secreted.  
CC -1- TISSUE SPECIFICITY: Expressed by the venom gland.  
CC -1- SIMILARITY: BELONGS TO THE SCORPION LEIUTOXIN FAMILY.

This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL Outstation at the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/ or send an email to license@isb-sib.ch).

EMBL; AF097408; AAF01253.1; --  
HSSP; Q9NJ7P; 1DU9.  
KW Toxin; Neurotoxin; Ionic channel inhibitor;  
KW Potassium channel inhibitor; Signal;  
FT SIGNAL 1 28  
FT CHAIN 29 56 NEUROTOXIN BMP03.  
FT DISULFID 31 47  
FT DISULFID 34 52  
FT DISULFID 38 54  
SQ SEQUENCE 56 AA; 6001 MW; 70953032007E8672 CRC64;

Query Match 39.5%; Score 30; DB 1; Length 56;  
Best Local Similarity 26.7%; Pred. No. 80;  
Matches 4; Conservative 8; Mismatches 3; Indels 0; Gaps 0;

QY 1 RYVDALNVLMMNNII 15  
DB 3 RLFTLVLLVLMNMV 17

RESULT 4  
YFCR\_PROVU STANDARD; PRT; 96 AA.

ID YFCR\_PROVU  
AC P20927; 1991 (Rel. 17, Created)  
DT 01-FEB-1991 (Rel. 17, Last sequence update)  
DE Mesobuthus martensii (Manchurian scorpion) (Buthus martensii).  
OC Eukaryota; Metazoa; Arthropoda; Chelicerata; Arachnida; Scorpiones;  
OC Buthoidea; Buthidae; Mesobuthus.  
NCBI\_TaxID=34649;

[1]  
SEQUENCE FROM N.A.  
RP TISSUE=Venom gland;  
RC MEDLINE=99313193; PubMed=10386622;  
RA Wu J.-J., Dai L., Lan Z.D., Chi C.-W.;  
RT "Genomic organization of three neurotoxins active on small conductance  
DE Neurotoxin Bmp03 precursor (potassium ion channel blocker P03P).  
OS Mesobuthus martensii (Manchurian scorpion) (Buthus martensii).  
OC Eukaryota; Metazoa; Arthropoda; Chelicerata; Arachnida; Scorpiones;  
OC Buthoidea; Buthidae; Mesobuthus.  
NCBI\_TaxID=34649;

[2]  
SEQUENCE FROM N.A.  
RP TISSUE=Venom gland;  
RC MEDLINE=99402983; PubMed=10471839;  
RA Zhu S.-Y., Li W.-X., Zeng X.-C., Jiang D.-H., Mao X., Liu H.;  
RT "Molecular cloning and sequencing of two 'short chain' and two 'long  
chain' K(+) channel-blocking peptides from the Chinese scorpion Buthus  
DE Martensii Karsch.";  
OS Martensii Karsch.";  
OC Buthoidea; Buthidae; Mesobuthus.  
NCBI\_TaxID=34649;

[3]  
FUNCTION: Blocks small conductance Calcium-activated potassium channels.  
CC -1- SUBCELLULAR LOCATION: Secreted.  
CC -1- TISSUE SPECIFICITY: Expressed by the venom gland.  
CC -1- SIMILARITY: BELONGS TO THE SCORPION LEIUTOXIN FAMILY.

This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL Outstation at the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/ or send an email to license@isb-sib.ch).

EMBL; AF097408; AAF01253.1; --  
HSSP; Q9NJ7P; 1DU9.  
KW Toxin; Neurotoxin; Ionic channel inhibitor;  
KW Potassium channel inhibitor; Signal;  
FT SIGNAL 1 28  
FT CHAIN 29 56 NEUROTOXIN BMP03.  
FT DISULFID 31 47  
FT DISULFID 34 52  
FT DISULFID 38 54  
SQ SEQUENCE 56 AA; 6001 MW; 70953032007E8672 CRC64;

Query Match 39.5%; Score 30; DB 1; Length 56;  
Best Local Similarity 26.7%; Pred. No. 80;  
Matches 4; Conservative 8; Mismatches 3; Indels 0; Gaps 0;

QY 1 RYVDALNVLMMNNII 15  
DB 3 RLFTLVLLVLMNMV 17

RESULT 4  
YFCR\_PROVU STANDARD; PRT; 96 AA.

ID YFCR\_PROVU  
AC P20927; 1991 (Rel. 17, Created)  
DT 01-FEB-1991 (Rel. 17, Last sequence update)  
DE Mesobuthus martensii (Manchurian scorpion) (Buthus martensii).<



```

RX MEDLINE=88004470; PubMed=3308458;
RA Cole S.T.;
RT "Nucleotide sequence and comparative analysis of the frd operon
RT encoding the fumarate reductase of Proteus vulgaris. Extensive
RT sequence divergence of the membrane anchors and absence of an
RT frd-linked ampC cephalosporinase gene.";
RL Eur. J. Biochem. 167:481-488(1987).
CC -!- SIMILARITY: BELONGS TO THE HUPF/HYPC FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; X06151; CAA29511.1; -.
DR PIR; S00119; S00119.
DR InterPro; IPR001109; HupF_HypC.
DR Pfam; PF01455; HupF_HypC; 1.
DR PIRSF; PIRSF005618; HupF_HypC; 1.
DR PRINTS; PR00445; HUPFHYPC.
DR ProDom; PD003112; HupF_HypC; 1.
DR TIGRFAMs; TIGR00074; hupF_hypc; 1.
DR PROSITE; PS01097; HUPF_HYPC; 1.
KW Hypothetical protein.
SQ SEQUENCE 96 AA; 10300 MW; EC094F1F37956BE3 CRC64;

Query Match 39.5%; Score 30; DB 1; Length 96;
Best Local Similarity 66.7%; Pred. No. 1.4e+02;
Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

OY 4 DALNVLMMAM 12
DB : || || ||
67 ETNALMMAM 75

RESULT 5
YOR YORA TTV1 STANDARD; PRT; 70 AA.
AC P19285;
DT 01-NOV-1990 (Rel. 16, Created)
DT 01-NOV-1990 (Rel. 16, Last sequence update)
DE Hypothetical 8.1 kDa protein.
OS Thermoproteus tenax virus 1 (strain KRA1) (TTV1).
OC Viruses; dsDNA viruses, no RNA stage; Lipothrrixviridae;
OC Lipothrrixvirus.
OX NCBI_TaxID=10480;
RN [1]
RP SEQUENCE FROM N.A.
RA Neumann H.;
RL Submitted (MAR-1989) to the EMBL/GenBank/DBJ databases.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; X14855; CAA32979.1; -.
DR Hypothetical protein.
SQ SEQUENCE 70 AA; 8125 MW; P88F0A8E46323EDC CRC64;

Query Match 38.2%; Score 29; DB 1; Length 70;
Best Local Similarity 38.5%; Pred. No. 1.5e+02;
Matches 5; Conservative 4; Mismatches 4; Indels 0; Gaps 0;

OY 2 VYDALNVLMMAM 14
DB || || : : ||

DB 32 VYVAITVTLNL 44

RESULT 6
NULM DASNO STANDARD; PRT; 98 AA.
ID O21333;
DT 15-JUL-1998 (Rel. 36, Created)
DT 15-JUL-1998 (Rel. 36, Last sequence update)
DE NADH-ubiquinone oxidoreductase chain 4L (EC 1.6.5.3).
GN MYND4L OR ND4L OR NADH4L.
OS Dasybus novemcinctus (Nine-banded armadillo).
OC Mitochondrion.
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Edentata; Dasypodidae; Dasybus.
OX NCBI_TaxID=9361;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=97357423; PubMed=9214749;
RA Aranson U., Gullberg A., Janke A.;
RT "Phylogenetic analyses of mitochondrial DNA suggest a sister group
RT relationship between Xenarthra (Edentata) and Ferungulataes.";
RL Mol. Biol. Evol. 14:762-768(1997).
CC -!- CATALYTIC ACTIVITY: NADH + ubiquinone = NAD(+) + ubiquinol.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; Y11832; CAA72524.1; -.
DR PIR; T11449; T11449.
DR InterPro; IPR001133; Oxidored_4L.
DR Pfam; PF00420; oxidored_q2; 1.
DR ProDom; PD000359; Oxidred4L; 1.
DR Oxidoreductase; NAD; Ubiquinone; Mitochondrion.
SQ SEQUENCE 98 AA; 10840 MW; D0FF9BC309048774 CRC64;

Query Match 38.2%; Score 29; DB 1; Length 98;
Best Local Similarity 55.6%; Pred. No. 2.2e+02;
Matches 5; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

OY 6 LNVLMAMNI 14
DB || : || : ||
6 LNIIMAFSI 14

RESULT 7
RETS_BOVIN STANDARD; PRT; 42 AA.
ID RET5_BOVIN
AC P82708;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Retinol-binding protein III, cellular (CRBP-III) (Fragment).
GN RBP5.
OS Bos taurus (Bovine).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidea;
OC Bovidae; Bovinae; Bos.
OX NCBI_TaxID=9913;
RN [1]
RP SEQUENCE, AND FUNCTION.
RX TISSUE=Kidney;
RX MEDLINE=21173623; PubMed=11274389;
RA Polli C., Calderone V., Ottonello S., Bolchi A., Zanotti G.,
RA Stoppini M., Berni R.;
RA "Identification, retinoid binding and X-ray analysis of a human
RT

```

[illegible]

```

RX MEDLINE=93190624; PubMed=8383392;
RA Shchelkunov S.N., Blinov V.M., Totmenin A.V., Marennikova S.S.,
RA Kolykhalov A.A., Frolov I.V., Chizhikov V.E., Gytarov V.V.,
RA Gashikov P.V., Belanov E.F., Belavin P.A., Resenchuk S.M.,
RA Andzhaparidze O.G., Sandakhchiev L.S.;
RT "Nucleotide sequence analysis of variola virus HindIII M, L, I genome
RT fragments.";
RL Virus Res. 27:25-35(1993).
RN [3]
RP COMPLETE GENOME.
RC STRAIN=India-1967 / Isolate Ind3;
RX MEDLINE=92202281; PubMed=8384129;
RA Shchelkunov S.N., Blinov V.M., Sandakhchiev L.S.;
RT "Genes of variola and vaccinia viruses necessary to overcome the host
RT protective mechanisms.";
RL FEBS Lett. 319:80-83(1993).
RN [4]
RP SEQUENCE FROM N.A.
RC STRAIN=Bangladesh-1975;
RX MEDLINE=94088747; PubMed=8264798;
RA Maseung R.F., Esposito J.J., Liu L., Qi J., Utterback T.R.,
RA Knight J.C., Aubin L., Yuran T.E., Parsons J.M., Loparev V.N.,
RA Selivanov N.A., Cavallaro K.F., Karlavage A.R., Mahy B.W.J.,
RA Venter C.J.;
RT "Potential virulence determinants in terminal regions of variola
RT smallpox virus genome.";
RL Nature 366:748-751(1993).
RN [5]
RP SEQUENCE FROM N.A.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; X67119; CAA47559.1; -.
DR EMBL; X69198; CAA49000.1; -.
DR EMBL; L22579; AAA60807.1; -.
DR PIR; C36843; C36843.
DR PIR; T28497; T28497.
DR InterPro; IPR006803; Pox_15.
DR Pfam; PF04713; Pox_15; 1.
SQ SEQUENCE 79 AA; 8778 MW; AD4E703B701B5C86 CRC64;

Query Match 36.8%; Score 28; DB 1; Length 79;
Best Local Similarity 54.5%; Pred. No. 2.7e+02;
Matches 6; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 4 DALNYLMAMNI 14
| | | |
DB 3 DAITVLTAI 13

RESULT 11
PSBE NEPOL
ID PSBE NEPOL STANDARD; PRT; 83 AA.
AC Q9TKY1;
DT 15-SEP-2003 (Rel. 42, Created)
DT 15-SEP-2003 (Rel. 42, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE Cytochrome b559 alpha subunit (PSII reaction center subunit V).
GN PSBE.
OS Nephroselmis olivacea.
OC Chloroplast.
OC Eukaryota; Viridiplantae; Chlorophyta; Prasinophyceae;
OC Chlorodendreales; Chlorodendraceae; Nephroselmis.
OX NCBI TaxID=31312;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=NIES-484;
RX MEDLINE=9398694; PubMed=10468594;
RA Turmel M., Otis C., Lemieux C.;

```

```

RT The complete chloroplast DNA sequence of the green alga Nephroselmis
RT olivacea: insights into the architecture of ancestral chloroplast
RT genomes.";
RL Proc. Natl. Acad. Sci. U.S.A. 96:10248-10253(1999).
CC -!- FUNCTION: This b-type cytochrome is tightly associated with the
CC reaction center of photosystem II and possibly is part of the
CC water-oxidation complex.
CC -!- SUBUNIT: Heterodimer of an alpha subunit and a beta subunit.
CC -!- SIMILARITY: Belongs to the pabe / pabf family.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; AF137379; AAD54835.1; -.
DR HAMAP; MF_00642; -.
DR InterPro; IPR006216; Cyt_b559.
DR InterPro; IPR006217; Cyt_b559_alpha.
DR Pfam; PF00283; cytochr_b559; 1.
DR Pfam; PF00284; cytochr_b559a; 1.
DR TIGRFAMs; TIGR01332; cyt_b559_alpha; 1.
DR PROSITE; PS00537; CYTOCHROME_B559; 1.
KW Photosystem II; Heme; Electron transport; Transmembrane; Chloroplast.
FT DOMAIN 1 18 STROMAL (POTENTIAL).
FT TRANSMEM 19 39 POTENTIAL.
FT DOMAIN 40 83 LUMENAL (POTENTIAL).
FT METAL 23 23 IRON (HEME AXIAL LIGAND) (BY SIMILARITY).
SQ SEQUENCE 83 AA; 9404 MW; B267F441B857D23 CRC64;

Query Match 36.8%; Score 28; DB 1; Length 83;
Best Local Similarity 33.3%; Pred. No. 2.8e+02;
Matches 5; Conservative 4; Mismatches 6; Indels 0; Gaps 0;

QY 2 VYDALNYLMAMNIIS 16
| | | | |
DB 66 ITRFNAQQMDILT 80

RESULT 12
ATP8 CANPA
ID ATP8 CANPA STANDARD; PRT; 48 AA.
AC P17345;
DT 01-AUG-1990 (Rel. 15, Created)
DT 01-AUG-1990 (Rel. 15, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE ATP synthase protein 8 (EC 3.6.3.14) (ATPase subunit 8) (A6L).
GN ATP8 OR AAP1.
OS Candida parapsilosis (Yeast).
OC Mitochondrion.
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; mitosporic Saccharomycetales; Candida.
OX NCBI TaxID=5480;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CBS 7154 / SP1;
RX MEDLINE=90332440; PubMed=2143015;
RA Guerin E., Velours J., Guerin M.;
RT "Cloning and sequencing of a fragment of the linear mitochondrial DNA
RT of the yeast Candida parapsilosis supporting genes encoding subunit 8
RT of F0 ATP synthase and a putative t-RNA(Pro).";
RL Nucleic Acids Res. 18:4267-4267(1990).
CC -!- FUNCTION: THIS IS ONE OF THE CHAINS OF THE NONENZYMATIC COMPONENT
CC (CF(0) SUBUNIT) OF THE MITOCHONDRIAL ATPASE COMPLEX.
CC -!- CATALYTIC ACTIVITY: ATP + H(2)O + H(+) (In) = ADP + phosphate +
CC H(+) (Out).
CC -!- SUBCELLULAR LOCATION: Membrane-bound.
CC -!- SIMILARITY: BELONGS TO THE ATPASE PROTEIN 8 FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration

```

15-JUL-1999 (Rel. 38, Created)  
15-JUL-1999 (Rel. 38, Last sequence update)  
16-OCT-2001 (Rel. 40, Last annotation update)  
Virulence-associated protein D.  
VAPD OR HI0450.  
Haemophilus influenzae.

```

OC Bacteria; Proteobacteria; Gammaproteobacteria; Pasteurellales;
OC Pasteurellaceae; Haemophilus.
OX NCBI_TaxID=727;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-Rd / KW20 / ATCC 51907;
RX MEDLINE=95350630; PubMed=7542800;
RA Fleischmann R.D., Adams M.D., White O., Clayton R.A., Kirkness E.F.,
RA Kerlavage A.R., Bult C.J., Tomb J.-F., Dougherty B.A., Merrick J.M.,
RA McKenney K., Sutton G., Fitzhugh W., Fields C.A., Gocayne J.D.,
RA Scott J.D., Shirley R., Liu L.-I., Glodek A., Kelley J.M.,
RA Weidman J.F., Phillips C.A., Spriggs T., Hedblom E., Cotton M.D.,
RA Usterback I.R., Hanna M.C., Nguyen D.T., Saudek D.M., Brandon R.C.,
RA Fine L.D., Fritchman J.L., Fuhrmann J.L., Geoghegan N.S.M.,
RA Gnehm C.L., McDonald L.A., Small K.V., Fraser C.M., Smith H.O.,
RA Venter J.C.;
RT "Whole-genome random sequencing and assembly of Haemophilus influenzae
RT Rd.";
RL Science 269:496-512 (1995).
RN [2]
RP SEQUENCE OF 76-80.
RX MEDLINE=20137488; PubMed=10675023;
RA Langen H., Takacs B., Evers S., Berndt P., Lahm H.W., Wipf B.,
RA Gray C., Fountoulakis M.;
RT "Two-dimensional map of the proteome of Haemophilus influenzae.";
RL Electrophoresis 21:411-429 (2000).
CC -!- SIMILARITY: BELONGS TO THE VAPD FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; U32728; AAC22108.1; -.
DR PIR; C64069; C64069.
DR TIGR; H10450; -.
DR Pfam; PF04605; VapD_N; 1.
KW Virulence; Complete proteome.
SQ SEQUENCE 91 AA; 10543 MW; 70B23CDE28E083E7 CRC64;

Query Match 35.5%; Score 27; DB 1; Length 91;
Best Local Similarity 41.7%; Pred. No. 4.8e+02;
Matches 5; Conservative 2; Mismatches 5; Indels 0; Gaps 0;

QY 4 DALNVLMAMNII 15
DB 50 DMANLFQAMNAL 61

RESULT 16
FLIE_AQUAE
ID FLIE_AQUAE STANDARD; PRT; 93 AA.
AC O67242;
DT 30-MAY-2000 (Rel. 39, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Flagellar hook-basal body complex protein flieE.
GN FLIE OR AQ_1182.1 OR AQ_1182A.
OS Aquifex aeolicus.
OC Bacteria; Aquificae; Aquificales; Aquificaceae; Aquifex.
OX NCBI_TaxID=63363;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=VF5;
RX MEDLINE=98196666; PubMed=95373320;
RA Deckert G., Warren P.V., Gaasterland T., Young W.G., Lenox A.L.,
RA Graham D.E., Overbeek R., Snead M.A., Keller M., Aujay M., Huber R.,
RA Feldman R.A., Short J.M., Olson G.J., Swanson R.V.;
RT "The complete genome of the hyperthermophilic bacterium Aquifex
aeolicus.";

```

```

RL Nature 392:353-358 (1998).
CC -!- SIMILARITY: BELONGS TO THE FLIE FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; AE000727; AAC07208.1; -.
DR PIR; G70401; G70401.
DR InterPro; IPR001624; FlieE.
DR Pfam; PF02049; FlieE; 1.
DR PRINTS; PR01006; FLGHOOFLIE.
DR TIGRPFAMs; TIGR00205; flieE; 1.
KW Flagella; Complete proteome.
SQ SEQUENCE 93 AA; 10764 MW; 6180AEAI08FB9C39 CRC64;

Query Match 35.5%; Score 27; DB 1; Length 93;
Best Local Similarity 35.7%; Pred. No. 4.9e+02;
Matches 5; Conservative 4; Mismatches 5; Indels 0; Gaps 0;

QY 1 RYDANVLMAMNII 14
DB 80 KLEAYNELMKQV 93

RESULT 17
VAPD_ACTAC
ID VAPD_ACTAC STANDARD; PRT; 95 AA.
AC Q52243;
DT 15-JUL-1999 (Rel. 38, Created)
DT 15-JUL-1999 (Rel. 38, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE Virulence-associated protein D homolog.
OS Actinobacillus actinomycetemcomitans (Haemophilus
OG Plasmid pVT736-1.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Pasteurellales;
OC Pasteurellaceae; Actinobacillus.
OX NCBI_TaxID=714;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=95362676; PubMed=7543479;
RA Galli D.M., Leblanc D.J.;
RT "Transcriptional analysis of rolling circle replicating plasmid
RT pVT736-1: evidence for replication control by antisense RNA.";
RL J. Bacteriol. 177:4474-4480 (1995).
CC -!- SIMILARITY: BELONGS TO THE VAPD FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; L24000; AAC37126.1; -.
DR Pfam; PF04605; VapD_N; 1.
KW Plasmid; Virulence.
SQ SEQUENCE 95 AA; 11190 MW; 3424348E815BF62A CRC64;

Query Match 35.5%; Score 27; DB 1; Length 95;
Best Local Similarity 62.5%; Pred. No. 5e+02;
Matches 5; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 2 VYDALNVLM 9
DB 57 VYKAINLL 64

```

bio-synthesis (By similarity)

-1- PATHWAY: De novo fatty acid biosynthesis.  
-1- PPM: 4'-phosphopantetheine is transferred from CoA to a specific serine of the apo-ACP-like protein (potential).  
-1- SIMILARITY: Contains 1 acyl carrier domain.

This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL Outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

EMBL; AE002149; AAP30918.1; --  
InterPro; IPR006163; Pp\_bind.  
InterPro; IPR006162; Ppantne\_attach.  
Pfam; PF00550; pp-binding; 1.  
PROSITE; PSS00075; ACP\_DOMAIN; 1.  
PROSITE; PS00012; PHOSPHOPANTETHEINE; 1.  
Hypothetical protein; Fatty acid biosynthesis; Phosphopantetheine; Complete proteome.  
BINDING 36 36 PHOSPHOPANTETHEINE (BY SIMILARITY).  
SEQUENCE 77 AA; 8750 MW; 3CDB655FFBFA968C6 CRC64;

Query Match 34.9%; Score 26.5; DB 1; Length 77;  
Best Local Similarity 38.1%; Pred. No. 4.9e+02;  
Matches 8; Conservative 2; Mismatches 2; Indels 9; Gaps 1;

QY 4 DALNV-----LWAMNII 15  
||| | | | | |  
DB 23 DNLNIELKSLGIDSLSAMNLI 43

RESULT 20  
YCK9\_CYAPA  
ID YCK9 CYAPA STANDARD; PRT; 48 AA.  
AC P48330;  
DT 01-FEB-1996 (Rel. 33, Created)  
DT 01-FEB-1996 (Rel. 33, Last sequence update)  
DT 16-OCT-2001 (Rel. 40, Last annotation update)  
DT Hypothetical 5.7 kDa protein in PSEB-PSEB intergenic region (ORF48).  
DE Cyanophora paradoxa.  
OS Cyanelle.  
OC Eukaryota; Glaucocystophyceae; Cyanophoraceae; Cyanophora.  
OX NCBI\_TaxID=2762;  
[1]\_TaxID=2762;  
SEQUENCE FROM N.A.  
RC STRAIN=UTEX LB 555 / Pringsheim;  
RA Stirewalt V.L., Michalowski C.B., Loeffelhardt W., Bohnert H.J.,  
RA Bryant D.A.;  
RT "Nucleotide sequence of the cyanelle DNA from Cyanophora paradoxa."  
RL Plant Mol. Biol. Rep. 13:327-332(1995).  
[2]  
SEQUENCE FROM N.A.  
RC STRAIN=UTEX LB 555 / Pringsheim;  
RA Loeffelhardt W., Stirewalt V.L., Michalowski C.B., Annarella M.,  
RA Farley J.Y., Schluter W.M., Chung S.J., Newmann-Spallart C.,  
RA Steiner J.M., Jakowitsch J., Bohnert H.J., Bryant D.A.;  
RT "The complete sequence of the cyanelle genome of Cyanophora paradoxa:  
ET the genetic complexity of a primitive plastid."  
RL (In) Schenk H.F.A., Herrmann R., Jeon K.W., Mueller N.E.,  
RL Schweinmuller W. (eds.);  
RL Eukaryotism and Symbiosis, pp.40-48, Springer-Verlag, Heidelberg  
[1997].

This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL Outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

```

RESULT 18
INULM_BALMU STANDARD; PRT; 98 AA.
ID AC 1
P41301;
01-FEB-1995 (Rel. 31, Created)
01-FEB-1995 (Rel. 31, Last sequence update)
01-NOV-1997 (Rel. 35, Last annotation update)
NADH-ubiquinone oxidoreductase chain 4L [EC 1.6.5.3].
MTND4L OR ND4L OR NADH4L.
Balaenoptera musculus (Blue whale).
Mitochondrion.
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Cetartiodactyla; Cetacea; Mysticeti;
Balaenopteridae; Balaenoptera.
NCBI_TaxID=9771;
[1]
SEQUENCE FROM N.A.
MEDLINE=94141932; PubMed=8308901;
Arnaason U., Gullberg A.;
"Comparison between the complete mtDNA sequences of the blue and the
fin whale, two species that can hybridize in nature.";
J. Mol. Evol. 37:312-322(1993).
-- CATALYTIC ACTIVITY: NADH + ubiquinone = NAD(+) + ubiquinol.
-----
This SWISS-PROT entry is copyright. It is produced through a collaboration
the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).
-----
CC EMBL; X72204; CAA51003.1; -.
CC PIR; S41828; S41828.
CC InterPro; IPR001133; Oxidored_4L.
CC InterPro; IPR003214; Oxidred4L.
CC Pfam; PF004420; Oxidored_42; 1.
CC ProDom; PD000359; Oxidred4L; 1.
CC Oxidoreductase; NAD; Ubiquinone; Mitochondrion.
CC SEQUENCE 98 AA; 10747 MW; 9F770651FE65ED1B CRC64;
-----
Query Match 35.5%; Score 27; DB 1; Length 98;
Best Local Similarity 55.6%; Pred. No. 5.2e+02;
Matches 5; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

Qy 6 LNVLMAMNI 14
:|||||::
Db 6 MNVLMAFSM 14

RESULT 19
ACPH_UREPA STANDARD; PRT; 77 AA.
ID AC 1
IC 09PPY4;
28-FEB-2003 (Rel. 41, Created)
28-FEB-2003 (Rel. 41, Last sequence update)
15-SEP-2003 (Rel. 42, Last annotation update)
DE Acyl carrier protein homolog (ACP).
GN U0506.
OS Ureaplasma parvum (Ureaplasma urealyticum biotype 1).
OC Bacteria; Firmicutes; Mollicutes; Mycoplasmataceae; Ureaplasma.
OX NCBI_TaxID=134821;
[1]
SEQUENCE FROM N.A.
RC STRAIN=Serovar 3;
RX MEDLINE=20500219; PubMed=11048724;
RA Glass J.I., Lefkowitz E.J., Glass J.S., Heiner C.R., Chen E.Y.,
RA Cassell G.H.;
RT "The complete sequence of the mucosal pathogen Ureaplasma
urealyticum.";
RL Nature 407:757-762 (2000).
-- FUNCTION: Carrier of the growing fatty acid chain in fatty acid

```

```

CC -----
DR EMBL; U30821; AA81282.1; -.
DR PIR; T06939; T06939.
KW Hypothetical protein; Cyanalle.
SQ SEQUENCE 48 AA; 5705 MW; 078F47EF024C26B5 CRC64;

Query Match 34.2%; Score 26; DB 1; Length 48;
Best Local Similarity 40.0%; Pred. NO. 3.6e+02;
Matches 6; Conservative 5; Mismatches 4; Indels 0; Gaps 0;

Oy 2 VYDALNVLMMNNIIS 16
Db 32 VFNNLSALIDKHIIS 46

RESULT 21
RL32 ECOLI STANDARD; PRT; 56 AA.
ID AC P02435;
DT 21-JUL-1996 (Rel. 01, Created)
DT 21-JUL-1996 (Rel. 01, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE 50S ribosomal protein L32.
GN RMPF OR B1089 OR T1728 OR ECS1467 OR STM1191 OR STY1230 OR T1729 OR
GN SF1093.
OS Escherichia coli.
OS Escherichia coli O157:H7.
OS Salmonella typhimurium.
OS Salmonella typhi. and
OS Shigella flexneri.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Escherichia.
OX NCBI_TaxID=562, 83334, 602, 601, 623;
RN [1]
RP SEQUENCE FROM N.A.
RC SPECIES=E.coli;
RX MEDLINE=90008815; PubMed=2477362;
RA Tanaka Y., Teujimura A., Fujita N., Isono S., Isono K.;
RT "Cloning and analysis of an Escherichia coli operon containing the
RT rpnF gene for ribosomal protein L32 and the gene for a 30-kilodalton
RT protein.";
RL J. Bacteriol. 171:5707-5712 (1989).
RN [2]
RP SEQUENCE FROM N.A.
RC SPECIES=E.coli; STRAIN=K12 / MG1655;
RX MEDLINE=97426617; PubMed=9278503;
RA Blattner F.R., Plunkett G. III, Bloch C.A., Perna N.T., Burland V.,
RA Riley M., Collado-Vides J., Glasner J.D., Rode C.K., Mayhew G.F.,
RA Gregor J., Davis N.W., Kirkpatrick H.A., Goeden M.A., Rose D.J.,
RA Mau B., Shao Y.;
RT "The complete genome sequence of Escherichia coli K-12.";
RL Science 277:1453-1474 (1997).
RN [3]
RP SEQUENCE FROM N.A.
RC SPECIES=E.coli; STRAIN=K12;
RX MEDLINE=97061202; PubMed=8905232;
RA Oshima T., Aiba H., Baba T., Fujita K., Hayashi K., Honjo A.,
RA Ikemoto K., Inada T., Itoh T., Kajihara M., Kanai K., Kashimoto K.,
RA Kimura S., Kitagawa M., Makino K., Masuda S., Miki T., Mizobuchi K.,
RA Mori H., Motomura K., Nakamura Y., Nishimoto H., Nishio Y., Saito N.,
RA Sempel G., Seki Y., Tagami H., Takemoto K., Wada C., Yamamoto Y.,
RA Yano M., Horiuchi T.;
RT "A 718-kb DNA sequence of the Escherichia coli K-12 genome
RT corresponding to the 12.7-28.0 min region on the linkage map.";
RL DNA Res. 3:137-155 (1996).
RN [4]
RP SEQUENCE FROM N.A.
RC SPECIES=E.coli; STRAIN=O157:H7 / BDL933 / ATCC 700927;
RX MEDLINE=21074935; PubMed=11206551;
RA Perna N.T., Plunkett G. III, Burland V., Mau B., Glasner J.D.,
RA Rose D.J., Mayhew G.F., Evans P.S., Gregor J., Kirkpatrick H.A.,
RA Posfai G., Hackett J., Klink S., Boutin A., Shao Y., Miller L.,
RA Grobbeck E.J., Davis N.W., Lim A., Dimalanta E.T., Potamousis K.,

```

```

RA Apodaca J., Anantharaman T.S., Lin J., Yen G., Schwartz D.C.,
RA Welch R.A., Blattner F.R.;
RT "Genome sequence of enterohaemorrhagic Escherichia coli O157:H7.";
RL Nature 409:529-533 (2001).
RN [5]
RP SEQUENCE FROM N.A.
RC SPECIES=E.coli; STRAIN=O157:H7 / RIMD 0509952;
RX MEDLINE=21152331; PubMed=11258796;
RA Hayashi T., Makino K., Ohnishi M., Kurokawa K., Ishii K., Yokoyama K.,
RA Han C.-G., Ohtsubo E., Nakayama K., Murata T., Tanaka M., Tobe T.,
RA Iida T., Takami H., Honda T., Sasaoka C., Ogasawara N., Yasunaga T.,
RA Kuhara S., Shiba T., Hattori M., Shinagawa H.;
RT "Complete genome sequence of enterohaemorrhagic Escherichia coli
RT O157:H7 and genomic comparison with a laboratory strain K-12.";
RL DNA Res. 8:11-22 (2001).
RN [6]
RP SEQUENCE.
RC SPECIES=E.coli; STRAIN=K;
RX MEDLINE=76119562; PubMed=765258;
RA Wittmann-Liebold B., Greuer B., Pannenbecker R.;
RT "The primary structure of protein L32 from the 50S subunit of
RT Escherichia coli ribosomes.";
RL Hoppe-Seyler's Z. Physiol. Chem. 356:1977-1979 (1975).
RN [7]
RP SEQUENCE.
RC SPECIES=E.coli; STRAIN=MRE-600;
RX MEDLINE=98317265; PubMed=9642179;
RA Zhang Y., Cronan J.E. Jr.;
RT "Transcriptional analysis of essential genes of the Escherichia coli
RT fatty acid biosynthesis gene cluster by functional replacement with
RT the analogous Salmonella typhimurium gene cluster.";
RL J. Bacteriol. 180:3295-3303 (1998).
RN [9]
RP SEQUENCE FROM N.A.
RC SPECIES=S.typhimurium; STRAIN=LT2;
RX MEDLINE=21534948; PubMed=11677609;
RA McClelland M., Sanderson K.E., Spieth J., Clifton S.W., Latreille P.,
RA Courtney L., Porwollik S., Ali J., Dante M., Du F., Hou S., Layman D.,
RA Leonard S., Nguyen C., Scott K., Holmes A., Grewal N., Mulvaney E.,
RA Ryan E., Sun H., Florea L., Miller W., Stoneking T., Nhan M.,
RA Waterston R., Wilson R.K.;
RT "Complete genome sequence of Salmonella enterica serovar Typhimurium
RT LT2.";
RL Nature 413:852-856 (2001).
RN [10]
RP SEQUENCE FROM N.A.
RC SPECIES=S.typhi; STRAIN=CT18;
RX MEDLINE=21534947; PubMed=11677608;
RA Parkhill J., Dougan G., James K.D., Thomson N.R., Pickard D., Wain J.,
RA Churcher C., Mungall K.L., Bentley S.D., Holden M.T.G., Sebahia M.,
RA Baker S., Bachev D., Brooks K., Chillingworth T., Connor P.,
RA Cronin A., Davis P., Davies R.M., Dowd L., White N., Farrar J.,
RA Feltwell T., Hamlin N., Haque A., Hien T.T., Holroyd S., Jagsels K.,
RA Krogh A., Larsen T.S., Leather S., Moule S., O'Gaora P., Parry C.,
RA Quail M., Rutherford K., Simmonds M., Skelton J., Stevens K.,
RA Whitehead S., Barrall B.G.;
RT "Complete genome sequence of a multiple drug resistant Salmonella
RT enterica serovar Typhi CT18.";
RL Nature 413:848-852 (2001).
RN [11]
RP SEQUENCE FROM N.A.
RC SPECIES=S.typhi; STRAIN=TY2 / ATCC 700931;
RX MEDLINE=22531367; PubMed=12644504;
RA Deng W., Liou S.-R., Plunkett G. III, Mayhew G.F., Rose D.J.,
RA Burland V., Kodoyianni V., Schwartz D.C., Blattner F.R.;
RT "Comparative genomics of Salmonella enterica serovar Typhi strains Ty2

```



RT and CT18.";

RL J. Bacteriol. 195:2330-2337(2003).

RN [12]

RP SEQUENCE FROM N.A.

RC SPECIES=S.flexneri; STRAIN=301 / Serotype 2a;

RX MEDLINE=2272406; PubMed=12384590;

RA Jin Q., Yuan Z., Xu J., Wang Y., Shen Y., Lu W., Wang J., Liu H.,

RA Yang J., Yang F., Zhang X., Zhang J., Yang G., Wu H., Qu D., Dong J.,

RA Sun L., Xue Y., Zhao A., Gao Y., Zhu J., Kan B., Ding K., Chen S.,

RA Cheng H., Yao Z., He B., Chen R., Ma D., Qiang B., Wen Y., Hou Y.,

RA Yu J.;

RT "Genome sequence of Shigella flexneri 2a: insights into pathogenicity

RT through comparison with genomes of Escherichia coli K12 and O157.";

RL Nucleic Acids Res. 30:4432-4441(2002).

RN [13]

RP MASS SPECTROMETRY.

RC SPECIES=E.coli; STRAIN=K12 / ATCC 25404;

RX MEDLINE=99196679; PubMed=10094780;

RA Arnold R.J., Reilly J.P.;

RT "Observation of Escherichia coli ribosomal proteins and their

RT posttranslational modifications by mass spectrometry.";

RL Anal. Biochem. 269:105-112(1999).

CC -1- MASS SPECTROMETRY: MW=6315.1; METHOD=MALDI.

CC -1- SIMILARITY: BELONGS TO THE L32P FAMILY OF RIBOSOMAL PROTEINS.

CC -----

CC This SWISS-PROT entry is copyright. It is produced through a collaboration

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -

CC the European Bioinformatics Institute. There are no restrictions on its

CC use by non-profit institutions as long as its content is in no way

CC modified and this statement is not removed. Usage by and for commercial

CC entities requires a license agreement. (See <http://www.isb-sib.ch/announce/>

CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

CC -----

DR EMBL; M29698; AAC24575.1; -.

DR EMBL; AE000209; AAC74173.1; -.

DR EMBL; D90744; BAA35897.1; -.

DR EMBL; AE005319; AAG55835.1; -.

DR EMBL; AF002555; BAB34890.1; -.

DR EMBL; AF044668; AAC38646.1; -.

DR EMBL; AE008751; AAL20120.1; -.

DR EMBL; AL627269; CAD08315.1; -.

DR EMBL; AE016839; AAO69353.1; -.

DR EMBL; AE015136; AAN42712.1; -.

DR PIR; C90812; C90812.

DR PIR; G85671; G85671.

DR PIR; JVO048; R5EC32.

DR ECO2DBASE; I011.7; 6TH EDITION.

DR EcoGene; EGI0890; rpmf.

DR StyGene; SGI0698; rpmf.

DR HAMAP; MF\_00340; -; 1.

DR InterPro; IPR005718; S32\_bact.org.

DR TIGRFAMs; TIGR01031; rpmf\_bact; 1.

KW Ribosomal protein; Complete proteome.

FT INIT MET 0 0

SQ SEQUENCE 56 AA; 6315 MW; B74A9510E78E4E94 CRC64;

Query Match 34.2%; Score 26; DB 1; Length 56;

Best Local Similarity 28.6%; Pred. No. 4.3e+02;

Matches 4; Conservative 6; Mismatches 4; Indels 0; Gaps 0;

QY 1 RVYDALNVLMMNNI 14

DB 16 RSHDALTAIVSLSV 29

RESULT 22

ATPL\_STRMU STANDARD; PRT; 67 AA.

ID ATPL\_STRMU

AC P95783;

DT 01-NOV-1997 (Rel. 35, Created)

DT 01-NOV-1997 (Rel. 35, Last sequence update)

DT 28-FEB-2003 (Rel. 41, Last annotation update)

DE ATP synthase C chain (BC 3.6.3.14) (Lipid-binding protein)

(Dicyclohexylcarbodiimide-binding protein).

ATPE OR SMU.1534.

OS Streptococcus mutans.

OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;

OC Streptococcus.

OX NCBI\_TaxID=1309;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=GS-5;

RX MEDLINE=97149283; PubMed=8996091;

RA Smith A.J., Quivey R.G., Faustoferri R.C.;

RT "Cloning and nucleotide sequence analysis of the Streptococcus mutans

RT membrane-bound, proton-translocating ATPase operon.";

RL Gene 183:87-96(1996).

RN [2]

RP SEQUENCE FROM N.A.

RC STRAIN=UAI59 / ATCC 700610 / Serotype C;

RX MEDLINE=22295063; PubMed=12397186;

RA Ajdic D., McShan W.M., McLaughlin R.E., Savic G., Chang J.,

RA Carson M.B., Primeaux C., Tian R., Kenton S., Jia H., Lin S., Qian Y.,

RA Li S., Zhu H., Najjar P., Lai H., White J., Roe B.A., Ferretti J.J.;

RT "Genome sequence of Streptococcus mutans UAI59, a cariogenic dental

RT pathogen.";

RL Proc. Natl. Acad. Sci. U.S.A. 99:14434-14439(2002).

CC -1- FUNCTION: THIS IS ONE OF THE THREE CHAINS OF THE NONENZYMATIC

CC COMPONENT (CF(0) SUBUNIT) OF THE ATPase COMPLEX.

CC -1- CATALYTIC ACTIVITY: ATP + H(2)O + H(+) (In) = ADP + phosphate +

CC H(+) (Out)

CC -1- SUBUNIT: F-TYPE ATPASES HAVE 2 COMPONENTS, CF(1) - THE CATALYTIC

CC CORE - AND CF(0) - THE MEMBRANE PROTON CHANNEL. CF(1) HAS FIVE

CC SUBUNITS: ALPHA(3), BETA(3), GAMMA(1), DELTA(1), EPSILON(1). CF(0)

CC HAS THREE MAIN SUBUNITS: A, B AND C.

CC -1- MISCELLANEOUS: DICYCLOHEXYLCARBODIIMIDE (DCCD) INHIBITS ATPASE.

CC -1- SIMILARITY: BELONGS TO THE ATPase C CHAIN FAMILY.

CC -----

CC This SWISS-PROT entry is copyright. It is produced through a collaboration

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -

CC the European Bioinformatics Institute. There are no restrictions on its

CC use by non-profit institutions as long as its content is in no way

CC modified and this statement is not removed. Usage by and for commercial

CC entities requires a license agreement. (See <http://www.isb-sib.ch/announce/>

CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

CC -----

DR EMBL; U31170; AAD13377.1; -.

DR EMBL; AE014985; AAN59184.1; -.

DR PIR; JCS735; JCS735.

DR InterPro; IPR002379; ATPase\_Caub.

DR InterPro; IPR000454; Sub\_ATPase\_Caub.

DR Pfam; PF00137; ATP-synt\_C; 1.

DR PRINTS; PR00124; ATPASEC.

DR PROSITE; PS00605; ATPASE\_C; 1.

KW Hydrogen ion transport; CF(0); Transmembrane; Lipid-binding;

KW Complete proteome.

FT BINDING 53 53 DICYCLOHEXYLCARBODIIMIDE (BY SIMILARITY).

SQ SEQUENCE 67 AA; 7060 MW; 56717BDA6A12FEDE CRC64;

Query Match 34.2%; Score 26; DB 1; Length 67;

Best Local Similarity 38.5%; Pred. No. 5.2e+02;

Matches 5; Conservative 2; Mismatches 6; Indels 0; Gaps 0;

QY 2 VYDALNVLMMNNI 14

DB 37 MYGKQLTLMIMGV 49

RESULT 23

YIFL\_ECOLI STANDARD; PRT; 67 AA.

ID YIFL\_ECOLI

AC P39166; Q8X3Y5;

DT 01-FEB-1995 (Rel. 31, Created)

DT 01-FEB-1995 (Rel. 31, Last sequence update)

DT 28-FEB-2003 (Rel. 41, Last annotation update)

DE Hypothetical lipoprotein yifL precursor.



GN Y1FL OR B3808.1 OR C4729 OR Z5325 OR ECS4737.1 OR SF3886.  
OS Escherichia coli,  
OS Escherichia coli O157:H7, and  
OS Shigella flexneri  
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;  
OC Enterobacteriaceae; Escherichia.  
OX NCBI\_TaxID=562, 217992, 83334, 623;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC SPECIES=E.coli; STRAIN=K12 / MG1655;  
RX MEDLINE=92358234; PubMed=1379743;  
RA Daniels D.L., Plunkett G. III, Burland V.D., Blattner F.R.;  
RT "Analysis of the Escherichia coli genome: DNA sequence of the region  
RT from 84.5 to 86.5 minutes.";  
RL Science 257:771-778 (1992).  
RN [2]  
RP SEQUENCE FROM N.A.  
RC SPECIES=E.coli; STRAIN=K12;  
RA Glaser P., Sismelro O., Danchin A.;  
RL Submitted (JUN-1992) to the EMBL/GenBank/DBJ databases.  
RN [3]  
RP SEQUENCE FROM N.A.  
RC SPECIES=E.coli; STRAIN=O6:H1 / CFT073 / ATCC 700928;  
RX MEDLINE=22388234; PubMed=12471157;  
RA Welch R.A., Burland V., Plunkett G. III, Redford P., Roesch P.,  
RA Raeko D., Buckles E.L., Liou S.-R., Boutin A., Hackett J., Stroud D.,  
RA Mayhew G.F., Rose D.J., Zhou S., Schwartz D.C., Perna N.T.,  
RA Mobley H.L.T., Donnenberg M.S., Blattner F.R.;  
RT "Extensive mosaic structure revealed by the complete genome sequence  
RT of uropathogenic Escherichia coli.";  
RL Proc. Natl. Acad. Sci. U.S.A. 99:17020-17024 (2002).  
RN [4]  
RP SEQUENCE FROM N.A.  
RC SPECIES=E.coli; STRAIN=O157:H7 / EDL933 / ATCC 700927;  
RX MEDLINE=21074935; PubMed=11206551;  
RA Perna N.T., Plunkett G. III, Burland V., Mau B., Glasner J.D.,  
RA Rose D.J., Mayhew G.F., Evans P.S., Gregor J., Kirkpatrick H.A.,  
RA Posfai G., Hackett J., Klink S., Boutin A., Shao Y., Miller L.,  
RA Grotbeck E.J., Davis N.W., Lim A., Dimalanta E.T., Potamousis K.,  
RA Apodaca J., Anantharaman T.S., Lin J., Yen G., Schwartz D.C.,  
RA Welch R.A., Blattner F.R.;  
RT "Genome sequence of enterohaemorrhagic Escherichia coli O157:H7.";  
RL Nature 409:529-533 (2001).  
RN [5]  
RP SEQUENCE FROM N.A.  
RC SPECIES=E.coli; STRAIN=O157:H7 / RIMD 0509952;  
RX MEDLINE=22272406; PubMed=12384590;  
RA Hayashi T., Makino K., Ohnishi M., Kurokawa K., Ishii K., Yokoyama K.,  
RA Han C.-G., Ohtsubo E., Nakayama K., Murata T., Tanaka M., Toke T.,  
RA Iida T., Takami H., Honda T., Sasakawa C., Ogasawara N., Yasunaga T.,  
RA Kubara S., Shiba T., Hattori M., Shinagawa H.;  
RT "Complete genome sequence of enterohaemorrhagic Escherichia coli  
RT O157:H7 and genomic comparison with a laboratory strain K-12.";  
RL DNA Res. 8:11-22 (2001).  
RN [6]  
RP SEQUENCE FROM N.A.  
RC SPECIES=S.flexneri; STRAIN=301 / Serotype 2a;  
RX MEDLINE=22272406; PubMed=12384590;  
RA Jin Q., Yuan Z., Xu J., Wang Y., Shen Y., Lu W., Wang J., Liu H.,  
RA Yang J., Yang F., Zhang X., Zhang J., Yang G., Wu H., Qu D., Dong J.,  
RA Sun L., Xue Y., Zhao A., Gao Y., Zhu J., Kan B., Ding K., Chen S.,  
RA Cheng H., Yao Z., He B., Chen R., Ma D., Qiang B., Wen Y., Hou Y.,  
RA Yu J.;  
RT "Genome sequence of Shigella flexneri 2a: insights into pathogenicity  
RT through comparison with genomes of Escherichia coli K12 and O157.";  
RL Nucleic Acids Res. 30:4432-4441 (2002).  
RN [7]  
RP IDENTIFICATION.  
RC SPECIES=E.coli;  
RX MEDLINE=95075659; PubMed=7984428;  
RA Borodovsky M., Rudd K.E., Koonin E.V.;  
RT "Intrinsic and extrinsic approaches for detecting genes in a

RT bacterial genome.";  
RL Nucleic Acids Res. 22:4756-4767 (1994).  
CC -!- SUBCELLULAR LOCATION: Attached to the membrane by a lipid anchor  
CC (Potential).  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement. (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
DR EMBL; M87049; -; NOT ANNOTATED\_CDS.  
DR EMBL; AE000457; -; NOT ANNOTATED\_CDS.  
DR EMBL; X66782; -; NOT ANNOTATED\_CDS.  
DR EMBL; AE016769; AAN83162.1; -;  
DR EMBL; AE005612; AAG59001.1; -;  
DR EMBL; AP002567; -; NOT ANNOTATED\_CDS.  
DR EMBL; AE015395; AAN45322.1; -;  
DR PIR; E86067; E86067.  
DR EcoGene; EG12353; YifL.  
DR PROSITE; PS00013; PROKAR LIPOPROTEIN; 1.  
KW Hypothetical protein; Membrane; Lipoprotein; Signal;  
KW Complete proteome.  
FT SIGNAL 1 19 POTENTIAL.  
FT CHAIN 20 67 HYPOTHETICAL LIPOPROTEIN YIFL.  
FT LIPID 20 20 N-ACYL DIGLYCERIDE (POTENTIAL).  
FT CONFLICT 60 67 DGPSQVNY -> MVHPF (IN REF. 2).  
SQ SEQUENCE 67 AA; 7177 MW; 4A52CF48A55A7F9 CRC64;  
  
Query Match 34.2%; Score 26; DB 1; Length 67;  
Best Local Similarity 45.5%; Pred. No. 5.2e+02;  
Matches 5; Conservative 3; Mismatches 3; Indels 0; Gaps 0;  
  
QY 2 VYDALNVLMM 12  
| : | | | :  
Db 4 VFKALTLLTL 14  
  
RESULT 24  
Y010\_TROWT  
ID Y010\_TROWT STANDARD; PRT; 69 AA.  
AC P59486;  
DT 15-SEP-2003 (Rel. 42, Created)  
DT 15-SEP-2003 (Rel. 42, Last sequence update)  
DT 15-SEP-2003 (Rel. 42, Last annotation update)  
DE Hypothetical UPF0233 membrane protein TW010/TW010.  
GN TW010 OR TW010.  
OS Tropheryma whipplei (strain Twist) (Whipple's bacillus), and  
OS Tropheryma whipplei (strain TW08/27) (Whipple's bacillus).  
OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;  
OC Micrococciaceae; Cellulomonadaceae; Tropheryma.  
OX NCBI\_TaxID=203267, 218496;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=Twist / Genotype 2A;  
RA Raoult D., Audic S., Robert C., Ogata H., Suhre K., Drancourt M.,  
RA Claverie J.-M.;  
RT "Tropheryma whipplei illustrates the diversity of gene loss patterns  
RT in small genome bacterial pathogens.";  
RL Submitted (JUL-2002) to the EMBL/GenBank/DBJ databases.  
RN [2]  
RP SEQUENCE FROM N.A.  
RC STRAIN=TW08/27;  
RX MEDLINE=22495039; PubMed=12606174;  
RA Bentley S.D., Mairwald M., Murphy L.D., Pallen M.J., Yeats C.A.,  
RA Dover L.G., Norbertczak H.T., Besra G.S., Quail M.A., Harris D.E.,  
RA von Herbay A., Goble A., Rutter S., Squares R., Squares S.,  
RA Barrell B.G., parkhill J., Relman D.A.;  
RT "Sequencing and analysis of the genome of the Whipple's disease  
RT bacterium Tropheryma whipplei.";  
RL Lancet 361:627-634 (2003).



```

OX NCBI_TaxID=554;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=SCRI193;
RX MEDLINE=94018629; PubMed=8412685;
RA Mulholland V., Hinton J.C.D., Sidebotham J., Toth I.K.,
RA Hyman L.J., Perombelon M.C.M., Reeves P.J., Salmood G.P.C.;
RT "A pleiotropic reduced virulence (Rvi-) mutant of Erwinia carotovora
RT subspecies atroseptica is defective in flagella assembly proteins
RT that are conserved in plant and animal bacterial pathogens.";
RL Mol. Microbiol. 9:343-356(1993).
CC -!- FUNCTION: ROLE IN FLAGELLAR BIOSYNTHESIS:
CC -!- SUBCELLULAR LOCATION: Integral membrane protein (Potential).
CC -!- SIMILARITY: BELONGS TO THE FLIQ/MOPD/SPAQ FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; X72969; CAAS1477.1; -.
DR PIR; S35277; S35277.
DR InterPro; IPR002191; Bac_export_3.
DR Pfam; PF01313; Bac_export_3; 1.
DR PRINTS; PR00952; TYPE31MQPROT.
DR TIGRFAMs; TIGR01402; fliQ; 1.
KW Flagella; Transmembrane.
FT TRANSMEM 16 38
FT TRANSMEM 55 74
FT TRANSMEM 89 AA; 9641 MW; 20EC2F84A09E063B CRC64;
SQ SEQUENCE 89 AA; 9641 MW; 20EC2F84A09E063B CRC64;
-----
Query Match 34.2%; Score 26; DB 1; Length 89;
Best Local Similarity 40.0%; Pred. No. 7.1e+02;
Matches 4; Conservative 4; Mismatches 2; Indels 0; Gaps 0;

Oy 3 YDALNVLMMAM 12
|:|:|:|:
Db 11 YEAMKVALAL 20

RESULT 27
TRP6 BOVIN
ID _TRP6 BOVIN STANDARD; PRT; 89 AA.
AC Q9MTW0;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Transient receptor potential channel 6 (Fragment).
GN TRPC6 OR TRP6.
OS Bos taurus (Bovine).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
OC Bovidae; Bovinae; Bos.
OX NCBI_TaxID=9913;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Brain;
RX MEDLINE=20379038; PubMed=10816590;
RA Philipp S., Trost C., Wernat J., Rautmann J., Himmerkus N.,
RA Schroth G., Kretz O., Nastainczyk W., Cavallie A., Hoth M.,
RA Flockerzi V.;
RT "TRP4 (CB1) protein is part of native calcium release-activated Ca2+-
RT like channels in adrenal cells.";
RL J. Biol. Chem. 275:23965-23972(2000).
CC -!- FUNCTION: THOUGHT TO FORM A RECEPTOR-ACTIVATED NON-SELECTIVE
CC CALCIUM PERMEANT CATION CHANNEL. PROBABLY IS OPERATED BY A
CC PHOSPHATIDYLINOSITOL SECOND MESSENGER SYSTEM ACTIVATED BY RECEPTOR
CC TYROSINE KINASES OR G-PROTEIN COUPLED RECEPTORS. ACTIVATED BY
CC DIACYLGLYCEROL (DAG) IN A MEMBRANE-DELIMITED FASHION,

```

```

CC INDEPENDENTLY OF PROTEIN KINASE C (BY SIMILARITY). SEEMS NOT TO BE
CC ACTIVATED BY INTRACELLULAR CALCIUM STORE DEPLETION (BY
CC SIMILARITY).
CC -!- SUBCELLULAR LOCATION: Integral membrane protein (Probable).
CC -!- SIMILARITY: BELONGS TO THE TRANSIENT RECEPTOR FAMILY. STRPC
CC SUBFAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; AJ271069; CAC01677.1; -.
DR InterPro; IPR002111; Cat_Channel_Trlp.
KW Ionic channel; Transmembrane; Ion transport; Calcium channel.
FT NON_TER 1 1
FT TRANSMEM 1 21
FT TRANSMEM 68 88
FT NON_TER 89 89
FT NON_TER 89 89
SQ SEQUENCE 89 AA; 10315 MW; 8E1D54F051F1E7ED CRC64;
-----
Query Match 34.2%; Score 26; DB 1; Length 89;
Best Local Similarity 28.6%; Pred. No. 7.1e+02;
Matches 4; Conservative 6; Mismatches 4; Indels 0; Gaps 0;

Oy 2 VYDALNVLMMAM 15
|:|:|:|:|:
Db 71 VYNTWVIVLNLML 84

RESULT 28
VDMA_BPT3
ID _VDMA_BPT3 STANDARD; PRT; 89 AA.
AC P10309;
DT 01-MAR-1989 (Rel. 10, Created)
DT 01-MAR-1989 (Rel. 10, Last sequence update)
DT 01-AUG-1992 (Rel. 23, Last annotation update)
DE DNA maturase A (DNA packaging protein A) (GP18).
GN 18.
OS Bacteriophage T3.
OC Viruses; dsDNA viruses, no RNA stage; Caudovirales; Podoviridae;
OC T7-like viruses.
OX NCBI_TaxID=10759;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=86209997; PubMed=3010556;
RA Yamada M., Fujisawa H., Kato H., Hamada K., Minagawa T.;
RT "Cloning and sequencing of the genetic right end of bacteriophage T3
RT DNA.";
RL Virology 151:350-361(1986).
RN [2]
RP ERRATUM.
RA Yamada M., Fujisawa H., Kato H., Hamada K., Minagawa T.;
RL Virology 154:246-246(1986).
CC -!- FUNCTION: DURING THE GROWTH OF THIS PHAGE, DNA IS SYNTHESIZED
CC AS CONCATAMERS. DURING DNA PACKAGING MATURE MONOMERS ARE CUT
CC FROM THE CONCATAMERS.
CC -!- SUBUNIT: GP18 AND GP19 ASSOCIATE WITH DNA AND PROHEAD.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; M14784; AAA92525.1; -.
DR PIR; C23476; JVBPT3.
KW DNA packaging.

```

```

SQ SEQUENCE 89 AA; 10033 MW; A131FE25FE1A8594 CRC64;
Query Match 34.2%; Score 26; DB 1; Length 89;
Best Local Similarity 40.0%; Pred. No. 7.1e+02;
Matches 4; Conservative 5; Mismatches 1; Indels 0; Gaps 0;

Qy 1 RYVDALNVLM 10
Db 35 QLYNAINKLL 44

RESULT 29
VDMA BPT7 STANDARD; PRT; 89 AA.
AC P03693;
DT 21-JUL-1986 (Rel. 01, Created)
DT 21-JUL-1986 (Rel. 01, Last sequence update)
DT 01-AUG-1992 (Rel. 23, Last annotation update)
DE DNA maturase A (DNA packaging protein A) (GP18).
GN 18.
OS Bacteriophage T7.
OC Viruses; dsDNA viruses, no RNA stage; Caudovirales; Podoviridae;
OC T7-like viruses.
OX NCBI_TaxID=10760;
RN [1]
SEQUENCE FROM N.A.
RX MEDLINE=83241725; PubMed=6864790;
RA Dunn J.J., Studier F.W.;
RT "Complete nucleotide sequence of bacteriophage T7 DNA and the
locations of T7 genetic elements."
RL J. Mol. Biol. 166:477-535(1983).
CC -!- FUNCTION: DURING THE GROWTH OF THIS PHAGE, DNA IS SYNTHESIZED
AS CONCATAMERS. DURING DNA PACKAGING MATURE MONOMERS ARE CUT
FROM THE CONCATAMERS.
CC -!- SUBUNIT: GP18 AND GP19 ASSOCIATE WITH DNA AND PROHEAD.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).
CC -----
DR EMBL; V01146; CAA24437.1; -.
DR PIR; A04315; JVBPA7.
DR DNA packaging.
SQ SEQUENCE 89 AA; 10145 MW; 344B03EC22B42145 CRC64;
Query Match 34.2%; Score 26; DB 1; Length 89;
Best Local Similarity 40.0%; Pred. No. 7.1e+02;
Matches 4; Conservative 5; Mismatches 1; Indels 0; Gaps 0;

Qy 1 RYVDALNVLM 10
Db 35 QLYNAINKLL 44

RESULT 30
EF1B AERPE STANDARD; PRT; 90 AA.
AC Q9Y974;
DT 30-MAY-2000 (Rel. 39, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Elongation factor 1-beta (EF-1-beta) (aEF-1beta).
GN EF1B OR APE2480.
OS Aeropyrum pernix.
OC Archaea; Crenarchaeota; Thermoprotei; Desulfurococcales;
OC Desulfurococaceae; Aeropyrum.
OX NCBI_TaxID=56636;
RN [1]
SEQUENCE FROM N.A.

```

```

RC STRAIN=K1;
RX MEDLINE=993110339; PubMed=10382966;
RA Kawarabayashi Y., Hino Y., Horikawa H., Yamazaki S., Haikawa Y.,
RA Jin-no K., Takahashi M., Sekine M., Baba S.-I., Ankai A., Kosugi H.,
RA Hosoyama A., Fukui S., Nagai Y., Nishijima K., Nakazawa H.,
RA Takamiya M., Masuda S., Funahashi T., Tanaka T., Kudoh Y.,
RA Yamazaki J., Kishida N., Oguchi A., Aoki K.-I., Kubota K.,
RA Nakamura Y., Nomura N., Sako Y., Kikuchi H.;
RT "Complete genome sequence of an aerobic hyper-thermophilic
crenarchaeon, Aeropyrum pernix K1."
RL DNA Res. 6:83-101(1999).
CC -!- FUNCTION: PROMOTES THE EXCHANGE OF GDP FOR GTP IN EF-1-ALPHA/GDP,
CC THUS ALLOWING THE REGENERATION OF EF-1-ALPHA/GTP THAT COULD THEN
CC BE USED TO FORM THE TERNARY COMPLEX EF-1-ALPHA/GTP/AATRNA (BY
CC SIMILARITY).
CC -!- SIMILARITY: BELONGS TO THE EF-1-BETA/EF-1-DELTA FAMILY.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).
CC -----
DR EMBL; AP000064; BAA81496.1; ALT_INIT.
DR HSP; O27734; IGH8.
DR HAMAP; MF 00043; -.
DR InterPro; IPR004542; aEF-1_beta.
DR InterPro; IPR001326; EF1_BD.
DR Pfam; PF00736; EF1BD; 1.
DR TIGRFAMs; TIGR00489; aEF-1_beta; 1.
DR Elongation factor; Protein biosynthesis; Complete proteome.
KW SEQUENCE 90 AA; 9885 MW; B7BC94013536F4 CRC64;
SQ
Query Match 34.2%; Score 26; DB 1; Length 90;
Best Local Similarity 55.6%; Pred. No. 7.2e+02;
Matches 5; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

Qy 5 ALNVLAMN 13
Db 50 ALKULVAMN 58

RESULT 31
NULM RHISA STANDARD; PRT; 91 AA.
AC O99826;
DT 30-MAY-2000 (Rel. 39, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE NADH-ubiquinone oxidoreductase chain 4L (EC 1.6.5.3).
GN NDA1.
OS Rhizopneustes sanguineus (Brown dog tick).
OG Mitochondrion.
OC Eukaryota; Metazoa; Arthropoda; Chelicerata; Arachnida; Acari;
OC Parasitiformes; Ixodida; Ixodidae; Rhinoceros.
OX NCBI_TaxID=34632;
RN [1]
SEQUENCE FROM N.A.
RX MEDLINE=99083443; PubMed=9866211;
RA Black W.C. IV, Roehrdanz R.L.;
RT "Mitochondrial gene order is not conserved in arthropods: prostriate
and metastriate tick mitochondrial genomes."
RL Mol. Biol. Evol. 15:1772-1785(1998).
CC -!- CATALYTIC ACTIVITY: NADH + ubiquinone = NAD(+) + ubiquinol.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).
CC -----

```

CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch).

DR EMBL; AF081829; AAD05527.1; -;  
DR PIR; T11163; T11163.  
DR InterPro; IPR001133; Oxidored4L.  
DR InterPro; IPR003214; Oxidred4L.  
DR Pfam; PF00420; oxidored\_2; 1.  
DR ProDom; PD000359; Oxidred4L; 1.  
KW Oxidoreductase; NAD; Ubiquinone; Mitochondrion.  
SQ SEQUENCE 91 AA; 10598 MW; D2E51EC5F1D056EE CRC64;

Query Match 34.2%; Score 26; DB 1; Length 91;  
Best Local Similarity 50.0%; Pred. No. 7.3e+02;  
Matches 4; Conservative 4; Mismatches 0; Indels 0; Gaps 0;

QY 8 VLMANNII 15  
DB 82 LMM5NNLI 89

## RESULT 32

ID IAPP CAVPO STANDARD; PRT; 92 AA.  
AC P12966;  
DT 01-OCT-1989 (Rel. 12, Created)  
DT 01-OCT-1989 (Rel. 12, Last sequence update)  
DT 28-FEB-2003 (Rel. 41, Last annotation update)  
DE Islet amyloid polypeptide precursor (Amylin).  
GN IAPP.  
OS Cavia porcellus (Guinea pig).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Rodentia; Hystricognathi; Caviidae; Cavia.  
OX NCBI\_TaxID=10141;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=89345542; PubMed=2668946;  
RA Nishi M., Chan S.-J., Negamatsu S., Bell G.I., Steiner D.F.;  
RT "Conservation of the sequence of islet amyloid polypeptide in five mammals is consistent with its putative role as an islet hormone.";  
RL Proc. Natl. Acad. Sci. U.S.A. 86:5738-5742(1989).  
CC -!- FUNCTION: SELECTIVELY INHIBITS INSULIN-STIMULATED GLUCOSE UTILIZATION AND GLYCOGEN DEPOSITION IN MUSCLE, WHILE NOT AFFECTING ADIPOCYTE GLUCOSE METABOLISM.  
CC -!- SUBCELLULAR LOCATION: Secreted.  
CC -!- SIMILARITY: BELONGS TO THE CALCITONIN FAMILY.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

CC EMBL; M25387; AAA37040.1; -;  
DR PIR; D33542; D33542.  
DR InterPro; IPR000443; Amylin.  
DR InterPro; IPR001693; Calcitonin-like.  
DR InterPro; IPR002163; Calcitonin B.  
DR Pfam; PF00214; Calc\_CGRP\_IAPP; 1.  
DR PRINTS; PR00817; CALCITONINB.  
DR PRINTS; PR00818; ISLETAMYLID.  
DR SMART; SM00113; CALCITONIN; 1.  
DR PROSITE; PS00258; CALCITONIN; 1.  
KW Hormone; Cleavage on pair of basic residues; Amidation; Amyloid; Signal.  
FT SIGNAL. 1 22 POTENTIAL.  
FT PROPEP 23 34  
FT PEPTIDE 37 73 ISLET AMYLOID POLYPEPTIDE.  
FT PROPEP 77 92  
FT DISULFID 38 43 BY SIMILARITY.  
FT MOD\_RES 73 73 AMIDATION (G-74 PROVIDE AMIDE GROUP).  
SQ SEQUENCE 92 AA; 9989 MW; 67F3629014BF3F9C CRC64;

Query Match 34.2%; Score 26; DB 1; Length 92;  
Best Local Similarity 58.3%; Pred. No. 7.4e+02;  
Matches 7; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY 5 ALNVLMANNIIIS 16  
DB 16 ALNELKATSIAS 27

## RESULT 33

ID YCHH\_ECOLI STANDARD; PRT; 92 AA.  
AC P31807;  
DT 01-JUL-1993 (Rel. 26, Created)  
DT 01-JUL-1993 (Rel. 26, Last sequence update)  
DT 28-FEB-2003 (Rel. 41, Last annotation update)  
DE Hypothetical protein ychH.  
GN YCHH OR B1205 OR C1663 OR Z1976 OR ECS1710.  
OS Escherichia coli,  
OS Escherichia coli O157:H7.  
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;  
OC Enterobacteriaceae; Escherichia.  
OX NCBI\_TaxID=562, 217992, 83334;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=K12;  
RX MEDLINE=92007806; PubMed=1833189;  
RA Garcia-Villagas M.R., de la Vega F.M., Galindo J.M., Segura M., Buckingham R.H., Guarneros G.;  
RT "Peptidyl-tRNA hydrolase is involved in lambda inhibition of host protein synthesis.";  
RL EMBO J. 10:3549-3555(1991).  
RN [2]  
RP REVISIONS.  
RC STRAIN=K12;  
RX MEDLINE=95129849; PubMed=7828865;  
RA Galindo J.M., de la Vega F.M., Guarneros G.;  
RT "Open reading frames flanking the peptidyl-tRNA hydrolase-encoding gene of Escherichia coli.";  
RL Gene 151:153-156(1994).  
RN [3]  
RP SEQUENCE FROM N.A.  
RC STRAIN=K12 / MG1655;  
RX MEDLINE=97426617; PubMed=9278503;  
RA Blattner F.R., Plunkett G. III, Bloch C.A., Perna N.T., Burland V., Riley M., Collado-Vides J., Glasner J.D., Rode C.K., Mayhew G.F., Gregor J., Davis N.W., Kirkpatrick H.A., Goeden M.A., Rose D.J., Mau B., Shao Y.;  
RT "The complete genome sequence of Escherichia coli K-12.";  
RL Science 277:1453-1474(1997).  
RN [4]  
RP SEQUENCE FROM N.A.  
RC STRAIN=K12;  
RX MEDLINE=97061202; PubMed=8905232;  
RA Oshima T., Aiba H., Baba T., Fujita K., Hayashi K., Honjo A., Ikemoto K., Inada T., Itoh T., Kajihara M., Kanai K., Kashimoto K., Kimura S., Kitagawa M., Makino K., Masuda S., Miki T., Mizobuchi K., Mori H., Motomura K., Nakamura Y., Nishimoto H., Nishio Y., Saito N., Sampei G., Seki Y., Tagami H., Takemoto K., Wada C., Yamamoto Y., Yano M., Horiuchi T.;  
RT "A 718-kb DNA sequence of the Escherichia coli K-12 genome corresponding to the 12.7-28.0 min region on the linkage map.";  
RL DNA Res. 3:137-155(1996).  
RN [5]  
RP SEQUENCE FROM N.A.  
RC STRAIN=O6:H1 / CFT073 / ATCC 700928;  
RX MEDLINE=22388234; PubMed=12471157;  
RA Welch R.A., Burland V., Plunkett G. III, Redford P., Roesch P., Rasko D., Buckles E.L., Liou S.-R., Boutin A., Hackett J., Stroud D., Mayhew G.F., Rose D.J., Zhou S., Schwartz D.C., Perna N.T., Mobley H.L.T., Donnenberg M.S., Blattner F.R.;

Tue Feb 17 11:55:53 2004

RT "Extensive mosaic structure revealed by the complete genome sequence  
of uropathogenic *Escherichia coli*.";  
Proc. Natl. Acad. Sci. U.S.A. 99:17020-17024 (2002).  
[6]  
SEQUENCE FROM N.A.  
RC STRAIN=O157:H7 / EDL933 / ATCC 700927;  
RX MEDLINE=21074935; PubMed=11206551;  
RA Perna N.T., Plunkett G. III, Burland V., Mau B., Glasner J.D.,  
Rose D.J., Mayhew G.F., Evans P.S., Gregor J., Kirkpatrick H.A.,  
Postfai G., Hackett J., Klink S., Boutin A., Shao Y., Miller L.,  
Grotbeck E.J., Davis N.W., Lim A., Dimalanta E.T., Potamousis K.,  
Apodaca J., Anantharaman T.S., Lin J., Yen G., Schwartz D.C.,  
Weich R.A., Blattner F.R.;  
RA "Genome sequence of enterohaemorrhagic *Escherichia coli* O157:H7";  
Nature 409:529-533 (2001).  
[7]  
SEQUENCE FROM N.A.  
RC STRAIN=O157:H7 / RMD 0509952;  
RX MEDLINE=21156231; PubMed=11258796;  
RA Hayashi T., Makino K., Ohnishi M., Kurokawa K., Ishii K., Yokoyama K.,  
Han C.-G., Ohtsubo E., Nakayama K., Murata T., Tanaka M., Tobe T.,  
Iida T., Takami H., Honda T., Sasakawa C., Ogasawara N., Yasunaga T.,  
Kuhara S., Shiba T., Hattori M., Shinagawa H.;  
RA "Complete genome sequence of enterohemorrhagic *Escherichia coli*  
O157:H7 and genomic comparison with a laboratory strain K-12";  
DNA Res. 8:11-22 (2001).  
-----  
This SWISS-PROT entry is copyright. It is produced through a collaboration  
between the Swiss Institute of Bioinformatics and the EMBL outstation -  
the European Bioinformatics Institute. There are no restrictions on its  
use by non-profit institutions as long as its content is in no way  
modified and this statement is not removed. Usage by and for commercial  
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
-----  
EMBL; X61941; CA943947.1; -;  
DR EMBL; AE000219; AAC74289.1; -;  
DR EMBL; D90756; BA336063.1; -;  
DR EMBL; AE016759; RAN80128.1; -;  
DR EMBL; AE005338; AAG56063.1; -;  
DR EMBL; AP002556; BAB35133.1; -;  
DR PIR; B64867; B64867.  
DR PIR; C85700; C85700.  
DR PIR; F90842; F90842.  
DR EcoGene; EG11533; ychH.  
KW Hypothetical protein; Complete proteome.  
SQ SEQUENCE 92 AA; 10486 MW; 8F75344F0C72525C CRC64;  
  
Query Match 34.2%; Score 26; DB 1; Length 92;  
Best Local Similarity 44.4%; Pred. No. 7.4e+02;  
Matches 4; Conservative 3; Mismatches 2; Indels 0; Gaps 0;  
  
QY 7 NVLMAMNII 15  
DB 11 NVLMGLGV 19  
||||| : : :  
RESULT 34  
C553 HELPJ  
ID \_C553 HELPJ STANDARD; PRT; 96 AA.  
AC Q9ZJ29;  
DT 28-FEB-2003 (Rel. 41, Created)  
DT 28-FEB-2003 (Rel. 41, Last sequence update)  
DT 28-FEB-2003 (Rel. 41, Last annotation update)  
DE Cytochrome C-553 precursor (C553).  
GN JHP1148.  
OS *Helicobacter pylori* J99 (Campylobacter *pylori* J99).  
OC Bacteria; Proteobacteria; Epsilonproteobacteria; Campylobacteriales;  
OC Helicobacteriaceae; Helicobacter.  
OX NCBI\_TaxID=85963;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=99120557; PubMed=9923682;

RA Alm R.A., Ling L.-S.L., Moir D.T., King B.L., Brown E.D., Doig P.C.,  
Smith D.R., Noonan B., Guild B.C., deJonge B.L., Carmel G.,  
Tummino P.J., Caruso A., Uria-Nickelsen M., Mills D.M., Ives C.,  
Gibson R., Merberg D., Mills S.D., Jiang Q., Taylor D.E., Vovis G.F.,  
Trust T.J.;  
RA "Genomic sequence comparison of two unrelated isolates of the human  
gastric pathogen *Helicobacter pylori*.";  
Nature 397:176-180 (1999).  
RT NATURE 397:176-180(1999)  
CC -!- FUNCTION: Natural electron acceptor for a formate dehydrogenase.  
CC -!- SUBCELLULAR LOCATION: Periplasmic.  
CC -!- PTM: BINDS ONE HEME GROUP PER MOLECULE.  
CC -!- SIMILARITY: BELONGS TO THE CLASS I CYTOCHROME C FAMILY.  
-----  
This SWISS-PROT entry is copyright. It is produced through a collaboration  
between the Swiss Institute of Bioinformatics and the EMBL outstation -  
the European Bioinformatics Institute. There are no restrictions on its  
use by non-profit institutions as long as its content is in no way  
modified and this statement is not removed. Usage by and for commercial  
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
-----  
EMBL; AE001542; AAD06721.1; -;  
DR PIR; F71843; F71843.  
DR HSP; P04032; ZDVH.  
DR InterPro; IPR003088; Cyt\_C1.  
DR InterPro; IPR002329; Cyt\_C1C.  
DR InterPro; IPR000345; Cyt\_C heme bind.  
DR Pfam; PF00034; cytochrome\_C\_1.  
DR PRINTS; PR00605; CYTOCHROME\_C1.  
DR PRODOM; PD004020; Cyt\_C\_bact; 1.  
DR PROSITE; PS00190; CYTOCHROME\_C; 1.  
KW Electron transport; Heme; Signal; Periplasmic; Complete proteome.  
FT SIGNAL 1 19  
FT CHAIN 20 96  
FT BINDING 29 29  
FT BINDING 32 32  
FT METAL 33 33  
FT METAL 73 73  
FT METAL 73 73  
FT METAL 73 73  
SQ SEQUENCE 96 AA; 10354 MW; 38607AE5D422AD82 CRC64;  
  
Query Match 34.2%; Score 26; DB 1; Length 96;  
Best Local Similarity 62.5%; Pred. No. 7.7e+02;  
Matches 5; Conservative 2; Mismatches 1; Indels 0; Gaps 0;  
  
QY 7 NVLMAMNI 14  
DB 15 NVLMATDV 22  
||||| : : :  
RESULT 35  
NULM BALPH  
ID NULM BALPH STANDARD; PRT; 98 AA.  
AC P24976;  
DT 01-MAR-1992 (Rel. 21, Created)  
DT 01-MAR-1992 (Rel. 21, Last sequence update)  
DT 16-OCT-2001 (Rel. 40, Last annotation update)  
DE NADH-ubiquinone oxidoreductase chain 4L (EC 1.6.5.3).  
GN MTND4L OR ND4L OR NADH4L.  
OS Balaenoptera physalus (finback whale) (Common rorqual).  
OC Mitochondrion.  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Cetartiodactyla; Cetacea; Mysticeti;  
OC Balaenopteridae; Balaenoptera.  
OX NCBI\_TaxID=9770;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX STRAIN=Isolate No. 27 / Anno 1987; TISSUE=Liver;  
RX MEDLINE=92139449; PubMed=1779436;  
RA Arnason U., Gullberg A., Widegren B.;  
RT "The complete nucleotide sequence of the mitochondrial DNA of the fin  
whale, *Balaenoptera physalus*.";

```

DR PROSITE; PS00050; RIBOSOMAL_L23; 1.
SQ Ribosomal protein; Chloroplast; rRNA-binding.
KW SEQUENCE 98 AA; 11944 MW; 7572EB84DA96D347 CRC64;

Query Match 33.6%; Score 25.5; DB 1; Length 98;
Best Local Similarity 47.1%; Pred. No. 9.7e+02;
Matches 8; Conservative 1; Mismatches 5; Indels 3; Gaps 1;

QY 2 VYDA---LNVLMAMNII 15
   |:|:|:|:|
Db 31 VFDVDKLNKLGQIKNII 47

RESULT 37
YCI12 CHLRE
ID YC12 CHLRE STANDARD; PRT; 33 AA.
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 15-JUL-1998 (Rel. 36, Last annotation update)
DE Hypothetical 3.3 kDa protein ycf12.
DE YCF12.
OS Chlamydomonas reinhardtii.
OG Chloroplast.
OC Eukaryota; Viridiplantae; Chlorophyta; Chlorophyceae; Volvocales;
OC Chlamydomonadaceae; Chlamydomonas.
NCBI_TaxID=3055;
[1] RN RN
RP SEQUENCE FROM N.A.
RC STRAIN=2137;
RT Khrabtukova I., Spreitzer R.J.;
RA "Nucleotide sequences of the chloroplast trnS-GCU and ycf12 genes of
RL Chlamydomonas reinhardtii.";
RL (In) Plant Gene Register PGR95-117.
CC -|- SIMILARITY: BELONGS TO THE YCF12 FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; U40346; AAA91171.1; -
DR FIR; T08018; T08018.
KW Chloroplast; Hypothetical protein.
SQ SEQUENCE 33 AA; 3332 MW; 022F633E416F349E CRC64;

Query Match 32.9%; Score 25; DB 1; Length 33;
Best Local Similarity 27.3%; Pred. No. 3.7e+02;
Matches 3; Conservative 6; Mismatches 2; Indels 0; Gaps 0;

QY 6 LNVLMAMNIIIS 16
   |:|:|:|:|
Db 1 MNIELALTVLS 11

RESULT 38
VG9_BPNF
ID VG9_BPNF STANDARD; PRT; 48 AA.
AC Q04395;
DT 15-DEC-1998 (Rel. 37, Created)
DT 15-DEC-1998 (Rel. 37, Last sequence update)
DT 15-DEC-1998 (Rel. 37, Last annotation update)
DE Tail protein (late protein Gp9) (Fragment).
DE GN 9.
OS Bacteriophage Nf.
OC Viruses; dsDNA viruses, no RNA stage; Caudovirales; Podoviridae;
OC phi-29-like viruses.
OC X
NCBI_TaxID=10753;
[1] RN RN
RP SEQUENCE FROM N.A.

```



Yoo S.K., Braithwaite D.K., Ito J.;  
"Head fiber protein gene of bacteriophage NF.";  
J. Gen. Appl. Microbiol. 36:365-368(1990).  
-----  
This SWISS-PROT entry is copyright. It is produced through a collaboration  
between the Swiss Institute of Bioinformatics and the EMBL outstation -  
the European Bioinformatics Institute. There are no restrictions on its  
use by non-profit institutions as long as its content is in no way  
modified and this statement is not removed. Usage by and for commercial  
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
-----  
EMBL; X52807; CAA36996.1; -;  
PIR; S11639; S11639.  
Late protein. 48  
NON\_TER 48  
SEQUENCE 48 AA; 5517 MW; 515CF4F4841414FC CRC64;  
Query Match 32.9%; Score 25; DB 1; Length 48;  
Best Local Similarity 33.3%; Pred. No. 5.5e+02;  
Matches 5; Conservative 4; Mismatches 6; Indels 0; Gaps 0;  
OY 1 RYVDALNVLMMNII 15  
Db 12 RIFSALRSPMIKIL 26  
-----  
RESULT 39  
LHAL ECTHL STANDARD; PRT; 52 AA.  
ID LHAL ECTHL  
AC P80102;  
DT 01-FEB-1994 (Rel. 28, Created)  
DT 01-FEB-1994 (Rel. 28, Last sequence update)  
DT 28-FEB-2003 (Rel. 41, Last annotation update)  
DE Light-harvesting protein B800/830/1020, alpha-1 chain (EHS-alpha-1)  
DE (Antenna pigment protein, alpha-1 chain) (Fragment).  
DE Ectothiorhodospira halochloris.  
OS Bacteria; Proteobacteria; Gammaproteobacteria; Chromatiales;  
OC Ectothiorhodospiraceae; Halorhodospira.  
OX NCBI\_TaxID=1052;  
RN [1]  
RP SEQUENCE  
RC STRAIN=DSM 1059;  
RX MEDLINE=92249336; PubMed=1577009;  
RA Wagner-Huber R., Brunisholz R.A., Bissig I., Frank G., Suter F.,  
Zuber H.;  
RA "The primary structure of the antenna polypeptides of  
RT Ectothiorhodospira halochloris and Ectothiorhodospira halophila. Four  
RT core-type antenna polypeptides in E. halochloris and E. halophila.";  
RL Eur. J. Biochem. 205:917-925(1992).  
CC -1- FUNCTION: ANTENNA COMPLEXES ARE LIGHT-HARVESTING SYSTEMS, WHICH  
CC TRANSFER THE EXCITATION ENERGY TO THE REACTION CENTERS.  
CC -1- SUBUNIT: THE CORE COMPLEX IS FORMED BY DIFFERENT ALPHA AND BETA  
CC CHAINS. BINDING BACTERIOCHLOROPHYLL MOLECULES, AND ARRANGED  
CC MOST PROBABLY IN TETRAMERIC STRUCTURES DISPOSED AROUND THE  
CC REACTION CENTER. THE NON-PIGMENTED GAMMA CHAINS MAY CONSTITUTE  
CC ADDITIONAL COMPONENTS.  
CC -1- SUBCELLULAR LOCATION: Type II membrane protein. Inner membrane.  
DR InterPro: IPR000066; Antenna\_a/b.  
DR InterPro: IPR002361; Antenna.  
DR Pfam: PF00556; LHC; 1.  
DR PROSITE; PS00968; ANTENNA\_COMP\_ALPHA; 1.  
KW Antenna complex; Light-harvesting polypeptide; Transmembrane;  
KW Magnesium; Bacteriochlorophyll; Inner membrane.  
FT DOMAIN 1 12 CYTOPLASMIC (POTENTIAL).  
FT TRANSMEM 13 33 POTENTIAL.  
FT DOMAIN 34 >52 PERIPLASMIC (POTENTIAL).  
FT METAL 29 29 MAGNESIUM (BACTERIOCHLOROPHYLL AXIAL  
FT LIGAND) (POTENTIAL).  
FT NON\_TER 52 52  
FT SEQUENCE 52 AA; 6191 MW; 85EID99B23CE6E0F CRC64;  
SQ  
Query Match 32.9%; Score 25; DB 1; Length 52;

Best Local Similarity 35.7%; Pred. No. 6e+02;  
Matches 5; Conservative 4; Mismatches 5; Indels 0; Gaps 0;  
OY 1 RYVDALNVLMMNII 14  
Db 6 KVFDPRLIATAI 19  
-----  
RESULT 40  
Y546 METJA STANDARD; PRT; 56 AA.  
ID Y546 METJA  
AC Q57966;  
DT 01-NOV-1997 (Rel. 35, Created)  
DT 01-NOV-1997 (Rel. 35, Last sequence update)  
DT 16-OCT-2001 (Rel. 40, Last annotation update)  
DE Hypothetical protein MJ0546.  
GN MJ0546.  
OS Methanococcus jannaschii.  
OC Archaea; Euryarchaeota; Methanococci; Methanococcales;  
OC Methanocaldococcaceae; Methanocaldococcus.  
OX NCBI\_TaxID=2190;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=JAL-1 / DSM 2661 / ATCC 43067;  
RX MEDLINE=96337999; PubMed=8688087;  
RA Bult C.J., White O., Olsen G.J., Zhou L., Fleischmann R.D.,  
RA Sutton G.G., Blake J.A., Fitzgerald L.M., Clayton R.A., Gocayne J.D.,  
RA Kerlavage A.R., Dougherty B.A., Tomb J.-F., Adams M.D., Reich C.I.,  
RA Overbeek R., Kirkness E.F., Weinstock K.G., Merrick J.M., Glodek A.,  
RA Scott J.L., Geoghagen N.S.M., Weidman J.P., Fuhrmann J.L., Nguyen D.,  
RA Utterback T.R., Kelley J.M., Peterson J.D., Sadow P.W., Hanna M.C.,  
RA Cotton M.D., Roberts K.M., Hurst M.A., Kaine B.P., Borodovsky M.,  
RA Klenk H.-P., Fraser C.M., Smith H.O., Woese C.R., Venter J.C.;  
RA "Complete genome sequence of the methanogenic archaeon, Methanococcus  
RT jannaschii".  
RL Science 273:1058-1073(1996).  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
CC EMBL; U67504; AAB98548.1; -;  
CC PIR; B64368; B64368.  
CC TIGR; MJ0546; -;  
KW Hypothetical protein; Transmembrane; Complete proteome.  
FT TRANSMEM 2 22 POTENTIAL.  
FT DOMAIN 27 47 GLU-RICH.  
SQ SEQUENCE 56 AA; 6460 MW; 66CBFC395548BA5D CRC64;  
Query Match 32.9%; Score 25; DB 1; Length 56;  
Best Local Similarity 30.0%; Pred. No. 6.5e+02;  
Matches 3; Conservative 5; Mismatches 2; Indels 0; Gaps 0;  
OY 6 LNVLMNMNII 15  
Db 13 LNIILGIKVI 22  
-----  
RESULT 41  
SCP1 MESMA STANDARD; PRT; 57 AA.  
ID SCP1 MESMA  
AC Q9U8D2; Q9U522;  
DT 28-FEB-2003 (Rel. 41, Created)  
DT 28-FEB-2003 (Rel. 41, Last sequence update)  
DT 28-FEB-2003 (Rel. 41, Last annotation update)  
DT 28-FEB-2003 (Rel. 41, Last annotation update)  
DE Neurotoxin BmP01 precursor (Potassium ion channel blocker P01).  
OS Mesobuthus martensii (Manchurian scorpion) (Buthus martensii).  
OC Eukaryota; Metazoa; Arthropoda; Chelicerata; Arachnida; Scorpiones;  
OC Buthoidea; Buthidae; Mesobuthus.





modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

EMBL; AE000242; AAC74537.1; --  
PIR; B64898; B64898.  
EcoGene; EG13778; YncH.  
HYPOTHETICAL protein; Complete proteome.  
SEQUENCE 70 AA; 8270 MW; 9DFEB75318F25C85 CRC64;

Query Match 32.9%; Score 25; DB 1; Length 70;  
Best Local Similarity 46.2%; Pred. No. 8.3e+02;  
Matches 6; Conservative 2; Mismatches 5; Indels 0; Gaps 0;

QY 2 VYDALNVLMAMNI 14  
Db 54 VGDALNINTNKNI 66

RESULT 44

DBHL\_RICCN STANDARD; PRT; 80 AA.  
AC Q92HL4;  
DT 28-FEB-2003 (Rel. 41, Created)  
DT 28-FEB-2003 (Rel. 41, Last sequence update)  
DT 28-FEB-2003 (Rel. 41, Last annotation update)  
DE DNA-binding protein HU-like.  
GN RC0757.

OS Rickettsia conorii.  
OC Bacteria; Proteobacteria; Alphaproteobacteria; Rickettsiales;  
OC Rickettsiaceae; Rickettsiae; Rickettsia.  
OX NCBI\_TaxID=781;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX STRAIN=Malish 7;  
RX MEDLINE=21442074; PubMed=11557893;  
RA Ogata H., Audic S., Renesto-Audiffren P., Fournier P.-E., Barbe V.,  
RA Samson D., Roux V., Cossart P., Weissenbach J., Claverie J.-M.,  
RA Raoult D.;  
RT "Mechanisms of evolution in Rickettsia conorii and R. prowazekii.";  
RL Science 293:2093-2098 (2001).  
CC -1- FUNCTION: THIS PROTEIN BELONGS TO THE HISTONE LIKE FAMILY OF  
CC PROKARYOTIC DNA-BINDING PROTEINS WHICH ARE CAPABLE OF WRAPPING  
CC DNA TO STABILIZE IT, AND PREVENT ITS DENATURATION UNDER EXTREME  
CC ENVIRONMENTAL CONDITIONS (BY SIMILARITY).  
CC -1- SIMILARITY: BELONGS TO THE BACTERIAL HISTONE-LIKE PROTEIN FAMILY.

-----  
This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

EMBL; AE008632; AL003295.1; --  
PIR; E97794; E97794.  
InterPro; IPR000119; Bac\_DNABind.  
Pfam; PF00216; Bac\_DNA\_Binding; 1.  
PROSITE; PS00045; HISTONE LIKE; FALSE NEG.  
KW DNA-binding; DNA condensation; Complete proteome.  
SQ SEQUENCE 80 AA; 9795 MW; 3B465A23F9C3DC2C CRC64;

Query Match 32.9%; Score 25; DB 1; Length 80;  
Best Local Similarity 44.4%; Pred. No. 9.6e+02;  
Matches 4; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 1 RYVDALNVL 9  
Db 10 KIHDKLNCL 18

RESULT 45

RANA\_RANPI STANDARD; PRT; 82 AA.  
AC P08950;  
DT 01-NOV-1988 (Rel. 09, Created)  
DT 01-NOV-1988 (Rel. 09, Last sequence update)  
DT 15-SEP-2003 (Rel. 42, Last annotation update)  
DE Ranatensin precursor.  
OS Rana pipiens (Northern leopard frog).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Amphibia; Batrachia; Anura; Neobatrachia; Ranoidea; Ranidae; Rana.  
OX NCBI\_TaxID=8404;

RN [1]  
RP SEQUENCE FROM N.A.  
RC TISSUE=Skin;  
RX MEDLINE=88330837; PubMed=2458345;  
RA Krane I.M., Naylor S.L., Helin-Davis D., Chin W.W., Spindel E.R.;  
RT "Molecular cloning of cDNAs encoding the human bombesin-like peptide  
RT encoding its amphibian homolog ranatensin.";  
RL J. Biol. Chem. 263:13317-13323 (1988).  
RN [2]  
RP SEQUENCE OF 48-58.

RA Nakajima T., Tanimura T., Pisano J.J.;  
RT "Isolation and structure of a new vasoactive polypeptide.";  
RL Fed. Proc. 29:282-282 (1970).  
CC -1- SUBCELLULAR LOCATION: Secreted.  
CC -1- TISSUE SPECIFICITY: Skin.  
CC -1- SIMILARITY: BELONGS TO THE BOMBESIN/NEUROMEDIN B/RANATENSIN FAMILY.

-----  
This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

EMBL; M21552; AAA49533.1; --  
PIR; B28945; B28945.  
InterPro; IPR000874; Bombesin.  
Pfam; PF02044; Bombesin; 1.  
PROSITE; PS00257; BOMBESIN; 1.  
KW Amphibian defense peptide; Bombesin family; Amidation;  
KW Cleavage on pair of basic residues; Signal.  
FT SIGNAL 1 27  
FT PROPEP 28 47  
FT PEPTIDE 48 58  
FT PROPEP 62 82  
FT MOD\_RES 58 58  
SQ SEQUENCE 82 AA; 9159 MW; 09194FFD0E7436DE CRC64;  
AMIDATION (G-59 PROVIDE AMIDE GROUP).

Query Match 32.9%; Score 25; DB 1; Length 82;  
Best Local Similarity 38.5%; Pred. No. 9.9e+02;  
Matches 5; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

QY 4 DALNVLMAMNI 16  
Db 13 DFLTILLFSPIS 25

Search completed: February 17, 2004, 10:57:04  
Job time : 6.06931 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: February 17, 2004, 10:50:13 ; Search time 9.66337 Seconds  
(without alignments)  
159.230 Million cell updates/sec

Title: US-09-900-147-5

Perfect score: 76

Sequence: 1 RYDALNVLMMNNIIS 16

Scoring table:

BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283308 seqs, 96168682 residues

Total number of hits satisfying chosen parameters: 37673

Minimum DB seq length: 0

Maximum DB seq length: 100

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

PIR 76.\*

1: pir1.\*

2: pir2.\*

3: pir3.\*

4: pir4.\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	35	46.1	60	JC1386	homeotic protein E
2	34	44.7	74	D89831	hypothetical prote
3	33	43.4	57	PX0624	alpha-internexin -
4	33	43.4	66	F84283	hypothetical prote
5	33	43.4	82	T42312	hypothetical prote
6	32	42.1	66	A99954	hypothetical prote
7	32	42.1	67	T27124	hypothetical prote
8	32	42.1	71	B86477	protein F1504_31
9	31	40.8	35	D83924	hypothetical prote
10	31	40.8	68	A12346	hypothetical prote
11	31	40.8	98	D89910	conserved hypothet
12	30	39.5	15	P80105	heat shock protein
13	30	39.5	29	E89904	hypothetical prote
14	30	39.5	34	E49410	t-complex polypept
15	30	39.5	35	A81346	hypothetical prote
16	30	39.5	59	D83797	hypothetical prote
17	30	39.5	60	A84215	hypothetical prote
18	30	39.5	72	B82699	hypothetical prote
19	30	39.5	82	T30649	hypothetical prote
20	30	39.5	94	I48172	glial fibrillary a
21	30	39.5	96	S00119	hypothetical prote
22	30	39.5	98	T17143	NADH2 dehydrogenas
23	30	39.5	98	T17174	NADH2 dehydrogenas
24	30	39.5	98	T17177	NADH2 dehydrogenas
25	30	39.5	98	T17162	NADH2 dehydrogenas
26	30	39.5	99	F90284	hypothetical prote
27	29.5	38.8	97	T40639	hypothetical prote
28	29	38.2	56	T03658	phosphoenolpyruvat
29	29	38.2	68	H85826	hypothetical prote

30 29 38.2 70 2 F89891 truncated transpos  
31 29 38.2 77 1 C69425 transcription regu  
32 29 38.2 78 2 T21426 hypothetical prote  
33 29 38.2 79 2 F97704 hypothetical prote  
34 29 38.2 82 2 E86742 protein-export pro  
35 29 38.2 82 2 E84290 hypothetical prote  
36 29 38.2 85 2 D97817 hypothetical prote  
37 29 38.2 98 2 T11449 NADH2 dehydrogenas  
38 29 38.2 98 2 JC5794 elongin C - fruit  
39 29 38.2 98 2 F89915 truncated transpos  
40 29 38.2 100 2 T38899 truncated transpos  
41 29 38.2 100 2 S53994 TodR protein - Pee  
42 28 36.8 30 2 S15650 NADH2 dehydrogenas  
43 28 36.8 45 2 C90001 hypothetical prote  
44 28 36.8 60 2 G69115 hypothetical prote  
45 28 36.8 63 2 AB2812 conserved hypothet

#### ALIGNMENTS

##### RESULT 1

JC1386

homeotic protein EgHbx1 - tapeworm (Echinococcus granulosus) (fragment)

C:Species: Echinococcus granulosus

C:Date: 30-Sep-1993 #sequence\_revision 30-Sep-1993 #text\_change 17-Oct-1997

C:Accession: JC1386

R:Oliver, G.; Vispo, M.; Mailhos, A.; Martinez, C.; Sosa-Pineda, B.; Fielitz, W.; Ehrlich

Gene 121, 337-342, 1992

A:Title: Homeoboxes in flatworms.

A:Reference number: JC1386; MUID:93077050; PMID:1359988

A:Accession: JC1386

A:Molecule type: DNA

A:Residues: 1-60 <OLI>

A:Cross-references: EMBL:X66817

C:Superfamily: unassigned homeobox proteins; homeobox homology

C:Keywords: DNA binding; homeobox; nucleus; transcription regulation

F:2-58/Domain: homeobox homology <HOX>

Query Match 46.1%; Score 35; DB 2; Length 60;

Best Local Similarity 38.5%; Pred. No. 21;

Matches 5; Conservative 6; Mismatches 2; Indels 0; Gaps 0;

QY 2 VYDALNVLMMNNI 14

DB 28 VYERLNLAISLNL 40

##### RESULT 2

D89831

hypothetical protein [imported] - Staphylococcus aureus (strain N315)

C:Species: Staphylococcus aureus

C:Date: 10-May-2001 #sequence\_revision 10-May-2001 #text\_change 22-Oct-2001

C:Accession: D89831

R:Kuroda, M.; Ohta, T.; Uchiyama, I.; Baba, T.; Yuzawa, H.; Kobayashi, I.; Cui, L.; Oguc

ma, A.; Mizutani-Ui, Y.; Kobayashi, N.; Sawano, T.; Inoue, R.; Kaito, C.; Sekimizu, K.;

C.; Shiba, T.; Hattori, M.; Ogasawara, N.; Hayashi, H.; Hiranatsu, K.

Lancet 357, 1225-1240, 2001

A:Title: Whole genome sequencing of methicillin-resistant Staphylococcus aureus.

A:Reference number: A89758; MUID:21311952; PMID:11418146

A:Accession: D89831

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-74 <KUR>

A:Cross-references: GB:BA000018; PID:gl3700510; PIDN:BA841807.1; GSPDB:GN00149

A:Experimental source: strain N315

C:Genetics:

A:Gene: SA0575

Query Match 44.7%; Score 34; DB 2; Length 74;

Best Local Similarity 50.0%; Pred. No. 40;

Matches 6; Conservative 4; Mismatches 2; Indels 0; Gaps 0;

T42312  
hypothetical protein - phage SPPI  
C:Species: phage SPPI  
C:Date: 03-Dec-1999 #sequence\_revision 03-Dec-1999 #text\_change 11-May-2000  
C:Accession: T42312  
R:Alonso, J.C.; Luder, G.; Stige, A.C.; Chai, S.; Weise, F.; Trautner, T.A.  
Gene 204, 201-212, 1997  
A:Title: The complete nucleotide sequence and functional organization of Bacillus subtilis phage SPPI  
A:Reference number: Z22137; MUID:98094274; PMID:9434185.  
A:Accession: T42312  
A:Status: preliminary; translated from GB/EMBL/DBJ

protein F1504.31 [imported] - Arabidopsis thaliana  
 C;Species: Arabidopsis thaliana (mouse-ear cress)  
 C;Date: 02-Mar-2001 #sequence\_revision 02-Mar-2001 #text\_change 31-Mar-2001  
 C;Accession: B86477  
 R;Theologis, A.; Ecker, J.R.; Palm, C.J.; Federpiel, N.A.; Kaul, S.; White, O.; Alonso,  
 Chin, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Conway, T.H.; Dewar, K.;  
 ansen, N.F.; Hughes, B.; Huizar, L.  
 Nature 408, 816-820, 2000  
 A;Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C.  
 C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Luros, J.S.; Maiti, R.; Marziali,  
 Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.  
 A;Authors: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallon,  
 ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.  
 A;Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.  
 A;Reference number: A86141; MUID:21016719; PMID:11130712  
 A;Accession: B86477  
 A;Status: preliminary  
 A;Molecule type: DNA  
 A;Residues: 1-71 <STO>  
 A;Cross-references: GB:AE005172; NID:98779365; PIDN:AAF79373.1; GSPDB:GN00141  
 C;Genetics:  
 A;Gene: F1504.31  
 A;Map position: 1

Query Match 42.1%; Score 32; DB 2; Length 71;  
 Best Local Similarity 33.3%; Pred. No. 88;  
 Matches 5; Conservative 5; Mismatches 5; Indels 0; Gaps 0;

Qy 1 RYDNLNVLMAANNII 15  
 :|||:|:|:|:  
 Db 35 RYDNLNVLMAANNII 49

RESULT 9  
 D83924  
 hypothetical protein BH2196 [imported] - Bacillus halodurans (strain C-125)  
 C;Species: Bacillus halodurans  
 C;Date: 01-Dec-2000 #sequence\_revision 01-Dec-2000 #text\_change 15-Jun-2001  
 C;Accession: D83924  
 R;Takami, H.; Nakasone, K.; Takaki, Y.; Maeno, G.; Sasaki, R.; Masui, N.; Fujii, F.; Hira  
 Nucleic Acids Res. 28, 4317-4331, 2000  
 A;Title: Complete genome sequence of the alkaliphilic bacterium Bacillus halodurans and  
 A;Reference number: A83650; MUID:20512582; PMID:11058132  
 A;Accession: D83924  
 A;Status: preliminary  
 A;Molecule type: DNA  
 A;Residues: 1-35 <STO>  
 A;Cross-references: GB:AP001514; GB:BA000004; NID:gl0174613; PIDN:BA05915.1; GSPDB:GN00  
 A;Experimental source: strain C-125  
 C;Genetics:  
 A;Gene: BH2196

Query Match 40.8%; Score 31; DB 2; Length 35;  
 Best Local Similarity 35.7%; Pred. No. 61;  
 Matches 5; Conservative 6; Mismatches 6; Indels 0; Gaps 0;

Qy 2 VYDNLNVLMAANNII 15  
 :|||:|:|:|:  
 Db 3 VYETLSIVVQNSI 16

RESULT 10  
 AI2346  
 hypothetical protein asl4328 [imported] - Nostoc sp. (strain PCC 7120)  
 C;Species: Nostoc sp. PCC 7120  
 A;Note: Nostoc sp. strain PCC 7120 is a synonym of Anabaena sp. strain PCC 7120  
 C;Date: 14-Dec-2001 #sequence\_revision 14-Dec-2001 #text\_change 09-Dec-2002  
 C;Accession: AI2346  
 R;Kaneko, T.; Nakamura, Y.; Wolk, C.P.; Kuritz, T.; Sasamoto, S.; Watanabe, A.; Iriguchi,  
 Nakazaki, N.; Shimpo, S.; Sugimoto, M.; Takazawa, M.; Yamada, M.; Yasuda, M.; Tabata, S  
 DNA Res. 8, 205-213, 2001  
 A;Title: Complete Genomic Sequence of the Filamentous Nitrogen-fixing Cyanobacterium Ana  
 A;Reference number: AB1807; MUID:21595285; PMID:11759840

A;Accession: AI2346  
 A;Status: preliminary  
 A;Molecule type: DNA  
 A;Residues: 1-68 <KUR>  
 A;Cross-references: GB:BA000019; PIDN:BA076027.1; PID:gl7133464; GSPDB:GN00179  
 A;Experimental source: strain PCC 7120  
 C;Genetics:  
 A;Gene: asl4328

Query Match 40.8%; Score 31; DB 2; Length 68;  
 Best Local Similarity 66.7%; Pred. No. 1.3e+02;  
 Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Qy 3 YDNLNVLMA 11  
 :|||:|:|:  
 Db 43 YEALNVLMA 51

RESULT 11  
 H89910  
 conserved hypothetical protein SAl186 [imported] - Staphylococcus aureus (strain N315)  
 C;Species: Staphylococcus aureus  
 C;Date: 10-May-2001 #sequence\_revision 10-May-2001 #text\_change 09-Dec-2002  
 C;Accession: H89910  
 R;Kuroda, M.; Ohta, T.; Uchiyama, I.; Baba, T.; Yuzawa, H.; Kobayashi, I.; Cui, L.; Oguc  
 ma, A.; Mizutani-Ui, Y.; Kobayashi, N.; Sawano, T.; Inoue, R.; Kaito, C.; Sekimizu, K.;  
 C.; Shiba, T.; Hattori, M.; Ogasawara, N.; Hayashi, H.; Hiramatsu, K.  
 Lancet 357, 1225-1240, 2001  
 A;Title: Whole genome sequencing of methicillin-resistant Staphylococcus aureus.  
 A;Reference number: A89758; MUID:21311952; PMID:11418146  
 A;Accession: H89910  
 A;Status: preliminary  
 A;Molecule type: DNA  
 A;Residues: 1-98 <KUR>  
 A;Cross-references: GB:BA000018; PID:gl3701149; PIDN:BA042444.1; GSPDB:GN00149  
 A;Experimental source: strain N315  
 C;Genetics:  
 A;Gene: SAl186  
 C;Superfamily: uncharacterized conserved protein

Query Match 40.8%; Score 31; DB 2; Length 98;  
 Best Local Similarity 66.7%; Pred. No. 1.9e+02;  
 Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Qy 3 YDNLNVLMA 11  
 :|||:|:|:  
 Db 62 YDDLNVVVA 70

RESULT 12  
 PA0105  
 heat shock protein dnaJ type - fungus (Fusarium sporotrichioides) (fragment)  
 C;Species: Fusarium sporotrichioides  
 C;Date: 20-Feb-1995 #sequence\_revision 20-Feb-1995 #text\_change 23-Mar-2001  
 C;Accession: PA0105  
 R;Chow, L.P.; Fukaya, N.; Sugiura, Y.; Ueno, Y.; Tabuchi, K.; Tsugita, A.  
 submitted to JIPID, October 1994  
 A;Description: Two dimensional polyacrylamide gel electrophoresis of Fusarium sporotrich  
 A;Reference number: PA0051  
 A;Accession: PA0105  
 A;Molecule type: protein  
 A;Residues: 1-15 <CHO>  
 C;Keywords: heat shock; stress-induced protein

Query Match 39.5%; Score 30; DB 2; Length 15;  
 Best Local Similarity 62.5%; Pred. No. 37;  
 Matches 5; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Qy 1 RYVDALNV 8  
 :|||:|:|:  
 Db 5 KLYDTLNV 12

## RESULT 13

E89904  
 hypothetical protein [imported] - Staphylococcus aureus (strain N315)  
 C:Species: Staphylococcus aureus  
 C>Date: 10-May-2001 #sequence\_revision 10-May-2001 #text\_change 22-Oct-2001  
 C:Accession: E89904  
 R:Kuroda, M.; Ohta, T.; Uchiyama, I.; Baba, T.; Yuzawa, H.; Kobayashi, I.; Cui, L.; Ogud  
 ma, A.; Mizutani-Ui, Y.; Kobayashi, N.; Sawano, T.; Inoue, R.; Kaito, C.; Sekimizu, K.;  
 C.; Shiba, T.; Hattori, M.; Ogasawara, N.; Hayashi, H.; Hiratsugu, K.  
 Lancet 357, 1225-1240, 2001  
 A>Title: Whole genome sequencing of methicillin-resistant Staphylococcus aureus.  
 A:Reference number: A89758; MUID:21311952; PMID:11418146  
 A:Accession: E89904  
 A>Status: preliminary  
 A:Molecule type: DNA  
 A:Residues: 1-29 <NR>  
 A:CROSS-references: GB:BA000018; PID:g13701098; PIDN:BA842393.1; GSPDB:GN00149  
 A:Experimental source: strain N315  
 C:Genetics:  
 A:Gene: SAS038

Query Match 39.5%; Score 30; DB 2; Length 29;  
 Best Local Similarity 55.6%; Pred. No. 76;  
 Matches 5; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

Qy 2 VYDALNVLM 10  
 :||| ||| :  
 Db 7 LYDTLNTLL 15

## RESULT 14

E49410  
 t-complex polypeptide 1 homolog (peak 2 fraction) - rabbit (fragments)  
 N:Alternate names: chaperonin homolog (peak 2)  
 C:Species: Oryctolagus cuniculus (domestic rabbit)  
 C>Date: 21-Sep-1994 #sequence\_revision 18-Nov-1994 #text\_change 04-Sep-1998  
 C:Accession: E49410  
 R:Rommelaere, H.; Van Troys, M.; Gao, Y.; Melki, R.; Cowan, N.J.; Vandekerckhove, J.; An  
 Proc. Natl. Acad. Sci. U.S.A. 90, 11975-11979, 1993  
 A>Title: Eukaryotic cytosolic chaperonin contains t-complex polypeptide 1 and seven rela  
 A:Reference number: A49410; MUID:94089752; PMID:7903455  
 A:Accession: E49410  
 A>Status: preliminary  
 A:Molecule type: protein  
 A:Residues: 1-34 <ROM>  
 A:Experimental source: reticulocyte  
 A>Note: sequence modified after extraction from NCBI backbone  
 A>Note: sequence extracted from NCBI backbone (NCBIP:141043)  
 C:Superfamily: molecular chaperone t-complex-type

Query Match 39.5%; Score 30; DB 2; Length 34;  
 Best Local Similarity 35.7%; Pred. No. 91;  
 Matches 5; Conservative 6; Mismatches 3; Indels 0; Gaps 0;

Qy 2 VYDALNVLMNMII 15  
 :||| :||| :  
 Db 3 LHDAMIVPRLLEII 16

## RESULT 15

A81346  
 hypothetical protein Cj0747 [imported] - Campylobacter jejuni (strain NCTC 11168)  
 C:Species: Campylobacter jejuni  
 C>Date: 31-Mar-2000 #sequence\_revision 31-Mar-2000 #text\_change 03-Jun-2002  
 C:Accession: A81346  
 R:Parkhill, J.; Wren, B.W.; Mungall, K.; Ketley, J.M.; Churcher, C.; Basham, D.; Chillin  
 C.W.; Quail, M.; Rajandream, M.A.; Rutherford, K.M.; VanVliet, A.; Whitehead, S.; Barrel  
 Nature 403, 665-668, 2000  
 A>Title: The genome sequence of the food-borne pathogen Campylobacter jejuni reveals hyp  
 A:Reference number: A81250; MUID:20150912; PMID:10688204  
 A:Accession: A81346  
 A>Status: preliminary  
 A:Molecule type: DNA

A;Residues: 1-35 <PAR>  
 A:CROSS-references: GB:AL139076; GB:AL111168; MID:G6968128; PIDN:CAB73017.1; PID:G696819  
 A:Experimental source: serotype O2, strain NCTC 11168  
 C:Genetics:  
 A:Gene: Cj0742; Cj0747

Query Match 39.5%; Score 30; DB 2; Length 35;  
 Best Local Similarity 25.9%; Pred. No. 93;  
 Matches 7; Conservative 5; Mismatches 1; Indels 14; Gaps 1;

Qy 2 VYDALNV-----LMNMNI 14  
 ||| :||| :  
 Db 4 VYDEINIAVNTCKGLVLSKRLALNL 30

## RESULT 16

D83797  
 hypothetical protein BHI180 [imported] - Bacillus halodurans (strain C-125)  
 C:Species: Bacillus halodurans  
 C>Date: 01-Dec-2000 #sequence\_revision 01-Dec-2000 #text\_change 15-Jun-2001  
 C:Accession: D83797  
 R:Takami, H.; Nakasone, K.; Takaki, Y.; Maeno, G.; Sasaki, R.; Masui, N.; Fujii, F.; Hira  
 Nucleic Acids Res. 28, 4317-4331, 2000  
 A>Title: Complete genome sequence of the alkaliphilic bacterium Bacillus halodurans and  
 A:Reference number: A83650; MUID:20512582; PMID:11058132  
 A:Accession: D83797  
 A>Status: preliminary  
 A:Molecule type: DNA  
 A:Residues: 1-59 <STO>  
 A:CROSS-references: GB:AP001511; GB:BA000004; MID:g10173727; PIDN:BA804899.1; GSPDB:GN00  
 A:Experimental source: strain C-125  
 C:Genetics:  
 A:Gene: BHI180

Query Match 39.5%; Score 30; DB 2; Length 59;  
 Best Local Similarity 31.2%; Pred. No. 1.7e+02;  
 Matches 5; Conservative 6; Mismatches 5; Indels 0; Gaps 0;

Qy 1 RVYDALNVLMNMNIIS 16  
 :||| :||| :  
 Db 36 KVFTINILMTILLTS 51

## RESULT 17

A84215  
 hypothetical protein Vng0568c [imported] - Halobacterium sp. NRC-1  
 C:Species: Halobacterium sp. NRC-1  
 C>Date: 02-Feb-2001 #sequence\_revision 02-Feb-2001 #text\_change 02-Feb-2001  
 C:Accession: A84215  
 R:Ng, W.V.; Kennedy, S.P.; Mahairas, G.G.; Berquist, B.; Pan, M.; Shukla, H.D.; Lasky, S  
 ; Leithausser, B.; Keller, K.; Cruz, R.; Danson, M.J.; Hough, D.W.; Maddocks, D.G.; Jabl  
 Jung, K.H.; Alam, M.; Freitas, T.  
 Proc. Natl. Acad. Sci. U.S.A. 97, 12176-12181, 2000  
 A:Authors: Hou, S.; Daniels, C.J.; Dennis, P.P.; Omer, A.D.; Ebhardt, H.; Lowe, T.M.; Li.  
 A>Title: Genome sequence of Halobacterium species NRC-1.  
 A:Reference number: A84160; MUID:20504483; PMID:11016950  
 A:Accession: A84215  
 A>Status: preliminary  
 A:Molecule type: DNA  
 A:Residues: 1-60 <STO>  
 A:CROSS-references: GB:AE004437; MID:g10580165; PIDN:AAG19085.1; GSPDB:GN00138  
 C:Genetics:  
 A:Gene: VNG0568C

Query Match 39.5%; Score 30; DB 2; Length 60;  
 Best Local Similarity 53.8%; Pred. No. 1.7e+02;  
 Matches 7; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

Qy 2 VYDALNVLMNMNI 14  
 ||| :||| :  
 Db 37 VYALINFLMSIAI 49

RESULT 18  
B82699  
hypothetical protein XF1305 [imported] - Xylella fastidiosa (strain 9a5c)  
C;Species: Xylella fastidiosa  
C;Date: 18-Aug-2000 #sequence\_revision 20-Aug-2000 #text\_change 20-Aug-2000  
C;Accession: B82699  
R;anonymous, The Xylella fastidiosa Consortium of the Organization for Nucleotide Sequencing  
Nature 406, 151-157, 2000  
A;Title: The genome sequence of the plant pathogen Xylella fastidiosa.  
A;Reference number: A82515; MUID:20365717; PMID:10910347  
A;Note: for a complete list of authors see reference number A59328 below  
A;Accession: B82699  
A;Status: preliminary  
A;Molecule type: DNA  
A;Residues: 1-72 <SIM>  
A;Cross-references: GB:AE003963; GB:AE003849; NID:g9106285; PIDN:AAF84114.1; GSPDB:GN001  
A;Experimental source: strain 9a5c  
R;Simpson, A.J.G.; Reinach, F.C.; Arruda, P.; Abreu, F.A.; Acencio, M.; Alvarenga, R.; Briones, M.R.S.; Bueno, M.R.P.; Camargo, A.A.; Camargo, L.E.A.; Carraro, D.M.; Carrer, H. as-Neto, E.; Docena, C.; El-Dorri, H.; Facincani, A.P.; Ferreira, A.J.S. submitted to GenBank, June 2000  
A;Authors: Ferreira, V.C.A.; Ferro, J.A.; Fraga, J.S.; Franca, S.C.; Franco, M.C.; Frohm J.D.; Junqueira, M.L.; Kemper, E.L.; Kitajima, J.P.; Krieger, J.E.; Kuramae, E.E.; Laigh chado, M.A.; Madeira, A.M.B.N.; Madeira, H.M.F.; Marino, C.L.; Marques, M.V.; Martins, E. A;Authors: Martins, E.M.F.; Matsukuma, A.Y.; Menck, C.F.M.; Miracca, E.C.; Miyaki, C.Y.; , F.G.; Nunes, L.R.; Oliveira, M.A.; de Oliveira, M.C.; de Oliveira, R.C.; Palmieri, D.A. Rodrigues, V.; Rosa, A.J. de M.; de Rosa Jr., V.E.; de Sa, R.G.; Santelli, R.V.; Sawasak M.; Tshako, M.H.; Vallada, H.; Van Sluys, M.A.; Verjovski-Almeida, S.; Vettore, A.L.; Z A;Reference number: A59328  
A;Contents: annotation  
C;Genetics:  
A;Gene: XF1305

Query Match 39.5%; Score 30; DB 2; Length 72;  
Best Local Similarity 66.7%; Pred. No. 2.1e+02;  
Matches 6; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 2 VYDALNVLM 10  
|||||  
Db 53 VYDAFNSFM 61

RESULT 19  
T30649  
hypothetical protein 47L - Molluscum contagiosum virus 1  
N;Alternate names: MC047L  
C;Species: Molluscum contagiosum virus 1  
C;Date: 05-Nov-1999 #sequence\_revision 05-Nov-1999 #text\_change 11-May-2000  
C;Accession: T30649  
R;Senkevich, T.G.; Bugert, J.J.; Sisler, J.R.; Koonin, E.V.; Darai, G.; Moss, B. Science 273, 813-816, 1996  
A;Title: Genome sequence of a human tumorigenic poxvirus: Prediction of specific host re A;Reference number: Z20876; MUID:96325459; PMID:8670425  
A;Accession: T30649  
A;Status: preliminary  
A;Molecule type: DNA  
A;Residues: 1-82 <SEN>  
A;Cross-references: EMBL:U60315; PIDN:AAC55175.1  
C;Genetics:  
A;Note: MC047L

Query Match 39.5%; Score 30; DB 2; Length 82;  
Best Local Similarity 50.0%; Pred. No. 2.4e+02;  
Matches 6; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

Qy 5 ALNVLMNNIIS 16  
|||||  
Db 13 ALTVLMMLWVS 24

RESULT 20  
148172

Query Match 39.5%; Score 30; DB 2; Length 94;  
Best Local Similarity 54.5%; Pred. No. 2.8e+02;  
Matches 6; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

Qy 4 DALNVLMAMNI 14  
|||||  
Db 18 DLLNVKLALDI 28

RESULT 21  
S00119  
hypothetical protein C - Proteus vulgaris  
C;Species: Proteus vulgaris  
C;Date: 31-Dec-1990 #sequence\_revision 31-Dec-1990 #text\_change 14-Apr-2003  
C;Accession: S00119  
R;Cole, S.T. Eur. J. Biochem. 167, 481-488, 1987  
A;Title: Nucleotide sequence and comparative analysis of the frd operon encoding the fun d-linked ampC cephalosporinase gene.  
A;Reference number: S00107; MUID:88004470; PMID:3308458  
A;Accession: S00119  
A;Molecule type: DNA  
A;Residues: 1-96 <COL>  
A;Cross-references: EMBL:X06151; NID:g45907; PIDN:CAA29511.1; PID:g45910  
C;Superfamily: [Nife]-hydrogenase maturation chapetone

Query Match 39.5%; Score 30; DB 2; Length 96;  
Best Local Similarity 66.7%; Pred. No. 2.8e+02;  
Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Qy 4 DALNVLMAM 12  
:|||||  
Db 67 ETLNLMAM 75

RESULT 22  
T17143  
NADH2 dehydrogenase (ubiquinone) (EC 1.6.5.3) chain 4L - Pygathrix bieti mitochondrion  
C;Species: Pygathrix bieti mitochondrion  
C;Date: 04-Mar-2000 #sequence\_revision 04-Mar-2000 #text\_change 03-Jun-2002  
C;Accession: T17143; T17165; T17168  
R;Wang, W.; Forstner, M.R.J.; Zhang, Y.P.; Lui, Z.M.; Wei, Y.; Huang, H.Q.; Hu, H.G.; Xi Int. J. Primatol. 18, 305-320, 1997  
A;Title: A phylogeny of Chinese leaf monkeys using mitochondrial ND3-ND4 gene sequences. A;Reference number: Z18709  
A;Accession: T17143  
A;Status: preliminary  
A;Molecule type: DNA  
A;Residues: 1-98 <WAN>  
A;Cross-references: EMBL:U92951; NID:g2290417; PID:g2290419; PIDN:AAD04678.1  
A;Experimental source: isolate DJ1  
A;Accession: T17165  
A;Status: preliminary  
A;Molecule type: DNA  
A;Residues: 1-98 <WA2>  
A;Cross-references: EMBL:U92957; NID:g2290441; PID:g2290443; PIDN:AAD08825.1  
A;Experimental source: isolate DJ4  
A;Accession: T17168

glial fibrillary acidic protein - golden hamster (fragment)  
C;Species: Mesocricetus auratus (golden hamster)  
C;Date: 02-Jul-1996 #sequence\_revision 02-Jul-1996 #text\_change 13-Aug-1999  
C;Accession: I48172  
R;Duguid, J.R.; Rohwer, R.G.; Seed, B. Proc. Natl. Acad. Sci. U.S.A. 85, 5738-5742, 1988  
A;Title: Isolation of cDNAs of scrapie-modulated RNAs by subtractive hybridization of a A;Reference number: I48171; MUID:88289792; PMID:2456582  
A;Accession: I48172  
A;Status: preliminary; translated from GB/EMBL/DBJ  
A;Molecule type: mRNA  
A;Residues: 1-94 <RES>  
A;Cross-references: GB:J03847; NID:g191446; PIDN:AAA37100.1; PID:g191447  
C;Superfamily: cytoskeletal keratin

A;Status: preliminary  
A;Molecule type: DNA  
A;Residues: 1-98 <W>3  
A;Cross-references: EMBL:U92958; NID:G2290445; PID:G2290447; PIDN:AAD08828.1  
A;Experimental source: isolate DJ5  
C;Superfamily: NADH dehydrogenase (ubiquinone) chain 4L  
C;Keywords: membrane-associated complex; mitochondrion; NAD; oxidative phosphorylation;

Query Match 39.5%; Score 30; DB 2; Length 98;  
Best Local Similarity 55.6%; Pred. No. 2.9e+02;  
Matches 5; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

Qy 6 LNVLMAMNI 14  
Db 43 MNTLMALNM 51

RESULT 23

T17174  
NADH2 dehydrogenase (ubiquinone) (EC 1.6.5.3) chain 4L - Pygathrix roxellana mitochondri  
C;Species: Pygathrix roxellana mitochondrion  
C;Date: 04-Mar-2000 #sequence\_revision 04-Mar-2000 #text\_change 03-Jun-2002  
C;Accession: T17174; T17180  
R;Wang, W.; Forstner, M.R.J.; Zhang, Y.P.; Lui, Z.M.; Wei, Y.; Huang, H.Q.; Hu, H.G.; Xi  
Int. J. Primatol. 18, 305-320, 1997  
A;Title: A phylogeny of Chinese leaf monkeys using mitochondrial ND3-ND4 gene sequences.  
A;Reference number: Z18709  
A;Accession: T17174  
A;Status: preliminary  
A;Molecule type: DNA  
A;Residues: 1-98 <W>3  
A;Cross-references: EMBL:U92961; NID:G2290457; PID:G2290459; PIDN:AAD04660.1  
A;Experimental source: isolate CUI  
A;Accession: T17180  
A;Status: preliminary  
A;Molecule type: DNA  
A;Residues: 1-98 <W>2  
A;Cross-references: EMBL:U92963; NID:G2290465; PID:G2290467; PIDN:AAD04666.1  
A;Experimental source: isolate CUI  
C;Superfamily: NADH dehydrogenase (ubiquinone) chain 4L  
C;Keywords: membrane-associated complex; mitochondrion; NAD; oxidative phosphorylation;

Query Match 39.5%; Score 30; DB 2; Length 98;  
Best Local Similarity 55.6%; Pred. No. 2.9e+02;  
Matches 5; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

Qy 6 LNVLMAMNI 14  
Db 43 MNTLMALNM 51

RESULT 24

T17177  
NADH2 dehydrogenase (ubiquinone) (EC 1.6.5.3) chain 4L [similarity] - Pygathrix roxellana  
C;Species: mitochondrion Pygathrix roxellana  
C;Date: 19-May-2000 #sequence\_revision 19-May-2000 #text\_change 03-Jun-2002  
C;Accession: T17177  
R;Wang, W.; Forstner, M.R.J.; Zhang, Y.P.; Lui, Z.M.; Wei, Y.; Huang, H.Q.; Hu, H.G.; Xi  
Int. J. Primatol. 18, 305-320, 1997  
A;Title: A phylogeny of Chinese leaf monkeys using mitochondrial ND3-ND4 gene sequences.  
A;Reference number: Z18709  
A;Accession: T17177  
A;Status: preliminary  
A;Molecule type: DNA  
A;Residues: 1-98 <W>3  
A;Cross-references: EMBL:U92962; NID:G2290461; PID:G2290463; PIDN:AAD04663.1  
C;Superfamily: NADH dehydrogenase (ubiquinone) chain 4L  
C;Keywords: membrane-associated complex; mitochondrion; NAD; oxidative phosphorylation;

Query Match 39.5%; Score 30; DB 2; Length 98;  
Best Local Similarity 55.6%; Pred. No. 2.9e+02;  
Matches 5; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

Qy 6 LNVLMAMNI 14  
Db 43 MNTLMALNM 51

RESULT 25

T17162  
NADH2 dehydrogenase (ubiquinone) (EC 1.6.5.3) chain 4L [similarity] - Pygathrix bieti mi  
C;Species: mitochondrion Pygathrix bieti  
C;Date: 19-May-2000 #sequence\_revision 19-May-2000 #text\_change 03-Jun-2002  
C;Accession: T17162  
R;Wang, W.; Forstner, M.R.J.; Zhang, Y.P.; Lui, Z.M.; Wei, Y.; Huang, H.Q.; Hu, H.G.; Xi  
Int. J. Primatol. 18, 305-320, 1997  
A;Title: A phylogeny of Chinese leaf monkeys using mitochondrial ND3-ND4 gene sequences.  
A;Reference number: Z18709  
A;Accession: T17162  
A;Status: preliminary  
A;Molecule type: DNA  
A;Residues: 1-98 <W>3  
A;Cross-references: EMBL:U92956; NID:G2290437; PID:G2290439; PIDN:AAD08822.1  
C;Superfamily: NADH dehydrogenase (ubiquinone) chain 4L  
C;Keywords: membrane-associated complex; mitochondrion; NAD; oxidative phosphorylation;

Query Match 39.5%; Score 30; DB 2; Length 98;  
Best Local Similarity 55.6%; Pred. No. 2.9e+02;  
Matches 5; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

Qy 6 LNVLMAMNI 14  
Db 43 MNTLMALNM 51

RESULT 26

F90284  
hypothetical protein SS01296 [imported] - Sulfolobus solfataricus  
C;Species: Sulfolobus solfataricus  
C;Date: 24-May-2001 #sequence\_revision 24-May-2001 #text\_change 24-May-2001  
C;Accession: F90284  
R;She, Q.; Singh, R.K.; Confalonieri, F.; Zivanovic, Y.; Allard, G.; Awayez, M.J.; Chan-  
Jong, I.; Jeffries, A.C.; Kozera, C.J.; Medina, N.; Peng, X.; Thi-Ngoc, H.P.; Redder, P.  
arrett, R.A.; Ragan, M.A.; Sensesen, C.W.; Van der Oost, J.  
submitted to GenBank, April 2001  
A;Description: Sulfolobus solfataricus complete genome.  
A;Reference number: A99139  
A;Accession: F90284  
A;Status: preliminary  
A;Molecule type: DNA  
A;Residues: 1-99 <KUR>  
A;Cross-references: GB:AE006641; NID:gl3814497; PIDN:AAK41533.1; GSPDB:GN00155  
C;Genetics:  
A;Gene: SS01296

Query Match 39.5%; Score 30; DB 2; Length 99;  
Best Local Similarity 36.4%; Pred. No. 2.9e+02;  
Matches 4; Conservative 6; Mismatches 1; Indels 0; Gaps 0;

Qy 6 LNVLMAMNIIS 16  
Db 74 INIISMNAIT 84

RESULT 27

T40639  
hypothetical protein SPBC685.08 - fission yeast (Schizosaccharomyces pombe)  
C;Species: Schizosaccharomyces pombe  
C;Date: 03-Dec-1999 #sequence\_revision 03-Dec-1999 #text\_change 03-Dec-1999  
C;Accession: T40639  
R;Seeger, K.; Harris, D.; Wood, V.; Rajandream, M.A.; Barrell, B.G.  
submitted to the EMBL Data Library, March 1999  
A;Reference number: Z21928  
A;Accession: T40639  
A;Status: preliminary; translated from GB/EMBL/DBJ  
A;Molecule type: DNA



A;Residues: 1-97 <SEE>  
A;Cross-references: EMBL:AL049474; PIDN:CAB39365.1; GSPDB:GN00067; SPDB:SPBC685.08  
A;Experimental source: strain 972h-; cosmid c685  
C;Genetics:  
A;Gene: SPDB:SPBC685.08  
A;Map position: 2  
A;Introns: 1/3

Query Match 38.8%; Score 29.5; DB 2; Length 97;  
Best Local Similarity 50.0%; Pred. No. 3.5e+02;  
Matches 7; Conservative 4; Mismatches 2; Indels 1; Gaps 1;

QY 2 VYDALNVL-MAMNI 14  
||||: |||: |||:  
DB 23 VYDAVKLLKLSALNL 36

RESULT 28  
T03658  
phosphoenolpyruvate carboxylase (EC 4.1.1.31) - maize (fragment)  
C;Species: Zea mays (maize)  
C;Date: 24-Mar-1999 #sequence\_revision 24-Mar-1999 #text\_change 22-Jun-1999  
C;Accession: T03658  
R;Yanagisawa, S.; Izui, K.; Yamaguchi, Y.; Shigesada, K.; Katsuki, H.  
FEBS Lett. 229, 107-110, 1988  
A;Title: Further analysis of cDNA clones for maize phosphoenolpyruvate carboxylase invol  
mRNA at multiple sites in vivo.  
A;Reference number: S00348; MUID:88152202; PMID:2894322  
A;Accession: T03658  
A;Status: Preliminary; translated from GB/EMBL/DBJ  
A;Molecule type: mRNA  
A;Residues: 1-56 <YAN>  
A;Cross-references: EMBL:X07171; NID:g22414; PIDN:CMA30159.1; PID:g929919  
C;Function:  
A;Description: catalyzes the irreversible beta-carboxylation of phosphoenolpyruvate by H  
A;Pathway: carbon dioxide fixation  
C;Superfamily: phosphoenolpyruvate carboxylase  
C;Keywords: carbon dioxide fixation; carbon-carbon lyase; carboxy-lyase; phosphoprotein

Query Match 38.2%; Score 29; DB 2; Length 56;  
Best Local Similarity 53.8%; Pred. No. 2.4e+02;  
Matches 7; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

QY 3 YDALNVL-MAMNII 15  
||||| |||: |||:  
DB 11 YDALLVDRFLNII 23

RESULT 29  
H05826  
hypothetical protein Z3179 [imported] - Escherichia coli (strain O157:H7, substrain EDL9  
C;Species: Escherichia coli  
C;Date: 16-Feb-2001 #sequence\_revision 16-Feb-2001 #text\_change 14-Sep-2001  
C;Accession: H05826  
R;Perna, N.T.; Plunkett III, G.; Burland, V.; Mau, B.; Glasner, J.D.; Rose, D.J.; Mayhew  
iller, L.; Grobeck, E.J.; Davis, N.W.; Lim, A.; Dimalanta, E.; Potamousis, K.; Apodaca,  
Nature 409, 529-533, 2001  
A;Title: Genome sequence of enterohemorrhagic Escherichia coli O157:H7.  
A;Reference number: A85480; MUID:21074935; PMID:11206551  
A;Accession: H05826  
A;Status: Preliminary  
A;Molecule type: DNA  
A;Residues: 1-68 <STO>  
A;Cross-references: GB:AE005174; NID:gi2516198; PIDN:AAG57076.1; GSPDB:GN00145; UNGP:Z31  
A;Experimental source: strain O157:H7, substrain EDL933  
C;Genetics:  
A;Gene: Z3179

Query Match 38.2%; Score 29; DB 2; Length 68;  
Best Local Similarity 54.5%; Pred. No. 3e+02;  
Matches 6; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 4 DALNVL-MAMNI 14

DB 13 DALRLVLQNI 23  
||||: |||: |||:  
RESULT 30  
F09891  
truncated transposase [imported] - Staphylococcus aureus (strain N315)  
C;Species: Staphylococcus aureus  
C;Date: 10-May-2001 #sequence\_revision 10-May-2001 #text\_change 22-Oct-2001  
C;Accession: F09891  
R;Kuroda, M.; Ohta, T.; Uchiyama, I.; Baba, T.; Yuzawa, H.; Kobayashi, I.; Cui, L.; Oguc  
ma, A.; Mizutani-Ui, Y.; Kobayashi, N.; Sawano, T.; Inoue, R.; Kaito, C.; Sekimizu, K.;  
C.; Shiba, T.; Hattori, M.; Ogasawara, N.; Hayashi, H.; Hiramatsu, K.  
Lancet 357, 1225-1240, 2001  
A;Title: Whole genome sequencing of methicillin-resistant Staphylococcus aureus.  
A;Reference number: A89758; MUID:21311952; PMID:11418146  
A;Accession: F09891  
A;Status: Preliminary  
A;Molecule type: DNA  
A;Residues: 1-70 <KUR>  
A;Cross-references: GB:BA000018; PID:gi3700994; PIDN:BAB42290.1; GSPDB:GN00149  
A;Experimental source: strain N315  
C;Genetics:  
A;Gene: truncated(tnp)

Query Match 38.2%; Score 29; DB 2; Length 70;  
Best Local Similarity 85.7%; Pred. No. 3e+02;  
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 8 VLMAMNI 14  
||||: |||:  
DB 22 VLMALNI 28

RESULT 31  
C69425  
transcription regulator asnC - Archaeoglobus fulgidus  
C;Species: Archaeoglobus fulgidus  
C;Date: 10-Sep-1999 #sequence\_revision 10-Sep-1999 #text\_change 21-Jul-2000  
C;Accession: C69425  
R;Klenk, H.P.; Clayton, R.A.; Tomb, J.F.; White, O.; Nelson, K.E.; Ketchum, K.A.; Dodson  
.; Fleischmann, R.D.; Quackenbush, J.; Lee, N.H.; Sutton, G.G.; Gill, S.; Kirkness, E.F.  
Glodek, A.; Zhou, L.; Overbeek, R.; Gocayne, J.D.; Weidman, J.F.; McDonald, L.  
Nature 390, 364-370, 1997  
A;Authors: Uterback, T.; Cotton, M.D.; Spriggs, T.; Artiach, P.; Kaine, B.P.; Sykes, S.  
Smith, H.O.; Woese, C.R.; Venter, J.C.  
A;Title: The complete genome sequence of the hyperthermophilic, sulfate-reducing archaeo  
A;Reference number: A69250; MUID:98049343; PMID:9389475  
A;Accession: C69425  
A;Status: Preliminary; nucleic acid sequence not shown; translation not shown  
A;Molecule type: DNA  
A;Residues: 1-77 <KLE>  
A;Cross-references: GB:AB001007; GB:AE000782; NID:g2689330; PIDN:AAB89844.1; PID:g264917  
C;Superfamily: Archaeoglobus fulgidus transcription regulator asnC

Query Match 38.2%; Score 29; DB 1; Length 77;  
Best Local Similarity 66.7%; Pred. No. 3.4e+02;  
Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 RVYDALNVL 9  
:||||| |  
DB 17 KVDYDALAAL 25

RESULT 32  
T21426  
hypothetical protein F26F2.4 - Caenorhabditis elegans  
C;Species: Caenorhabditis elegans  
C;Date: 15-Oct-1999 #sequence\_revision 15-Oct-1999 #text\_change 15-Oct-1999  
C;Accession: T21426  
R;Ainscough, R.  
submitted to the EMBL Data Library, November 1996  
A;Reference number: Z19420



A;Cross-references: EMBL:Y11832; NID:g2252500; PIDN:CAA72524.1; PID:g2252509  
C;Genetics:  
A;Genome: mitochondrion  
A;Genetic code: SGC1  
A;Note: NADH4L  
C;Superfamily: NADH dehydrogenase (ubiquinone) chain 4L  
C;Keywords: membrane-associated complex; mitochondrion; NAD; oxidative phosphorylation;

Query Match 38.2%; Score 29; DB 2; Length 98;  
Best Local Similarity 55.6%; Pred. No. 4.4e+02;  
Matches 5; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

Qy 6 LNLVLMANNI 14  
|||:|:|  
Db 6 LNIIMAFSI 14

RESULT 38  
JC5794  
elongin C - fruit fly (*Drosophila melanogaster*)  
C;Species: *Drosophila melanogaster*  
C;Date: 24-Jan-1998 #sequence\_revision 13-Mar-1998 #text\_change 17-Nov-2000  
C;Accession: JC5794  
R;Aso, T.; Conrad, M.N.  
Biochem. Biophys. Res. Commun. 241, 334-340, 1997  
A;Title: Molecular cloning of DNAs encoding the regulatory subunits of elongin from *Saccharomyces cerevisiae*  
A;Reference number: JC5792; MUID:98086352; PMID:9425272  
A;Accession: JC5794  
A;Status: nucleic acid sequence not shown  
A;Molecule type: mRNA  
A;Residues: 1-98 <ASO>  
A;Cross-references: DBJQ:AB007693; NID:g2780364; PIDN:BAA24287.1; PID:g2780365  
C;Comment: This protein binds to each other in the absence of elongin A to form a binary or induction of elongin A activity.  
C;Genetics:  
A;Cross-references: FlyBase:FBgn0023211  
C;Superfamily: elongin C

Query Match 38.2%; Score 29; DB 2; Length 98;  
Best Local Similarity 54.5%; Pred. No. 4.4e+02;  
Matches 6; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

Qy 5 ALNLVLMANNII 15  
|||:|:|  
Db 86 ALELLMAANFL 96

RESULT 39  
F89915  
truncated transposase [imported] - *Staphylococcus aureus* (strain N315)  
C;Species: *Staphylococcus aureus*  
C;Date: 10-May-2001 #sequence\_revision 10-May-2001 #text\_change 22-Oct-2001  
C;Accession: F89915  
R;Kuroda, M.; Ohta, T.; Uchiyama, I.; Baba, T.; Yuzawa, H.; Kobayashi, I.; Cui, L.; Ogud  
ma, A.; Mizutani-Ui, Y.; Kobayashi, N.; Sawano, T.; Inoue, R.; Kaito, C.; Sekimizu, K.;  
C.; Shiba, T.; Hattori, M.; Ogasawara, N.; Hayashi, H.; Hiramatsu, K.  
Lancet 357, 1225-1240, 2001  
A;Title: Whole genome sequencing of methicillin-resistant *Staphylococcus aureus*.  
A;Reference number: A89758; MUID:21311952; PMID:11418146  
A;Accession: F89915  
A;Status: preliminary  
A;Molecule type: DNA  
A;Residues: 1-98 <KUR>  
A;Cross-references: GB:BA000018; PID:gl3701187; PIDN:BA842482.1; GSPDB:GN00149  
A;Experimental source: strain N315  
C;Genetics:  
A;Gene: truncated-SA

Query Match 38.2%; Score 29; DB 2; Length 98;  
Best Local Similarity 85.7%; Pred. No. 4.4e+02;  
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 8 VLMMANNI 14

Db 59 VLMMANNI 65  
|||:|:|

RESULT 40  
T38899  
hypothetical protein SPAC513.04 - fission yeast (*Schizosaccharomyces pombe*)  
C;Species: *Schizosaccharomyces pombe*  
C;Date: 03-Dec-1999 #sequence\_revision 03-Dec-1999 #text\_change 04-Mar-2000  
C;Accession: T38899  
R;Seeger, K.; Harris, D.; McDougall, R.C.; Rajandream, M.A.; Barrell, B.G.  
submitted to the EMBL Data Library, October 1999  
A;Reference number: Z21816  
A;Accession: T38899  
A;Status: preliminary; translated from GB/EMBL/DBJ  
A;Molecule type: DNA  
A;Residues: 1-100 <SEE>  
A;Cross-references: EMBL:AL122032; PIDN:CAB59727.1; GSPDB:GN00066; SPDB:SPAC513.04  
A;Experimental source: strain 972h-; cosmid c513  
C;Genetics:  
A;Gene: SPDB:SPAC513.04  
A;Map position: 1  
C;Superfamily: *Schizosaccharomyces pombe* hypothetical protein SPAC513.04

Query Match 38.2%; Score 29; DB 2; Length 100;  
Best Local Similarity 33.3%; Pred. No. 4.5e+02;  
Matches 4; Conservative 5; Mismatches 3; Indels 0; Gaps 0;

Qy 1 RVDALNVLMM 12  
|||:|:|  
Db 12 RWSLNIICLM 23

RESULT 41  
SS3994  
TODR protein - *Pseudomonas putida*  
C;Species: *Pseudomonas putida*  
C;Date: 28-Oct-1996 #sequence\_revision 13-Mar-1997 #text\_change 08-Oct-1999  
C;Accession: SS3994  
R;Wang, Y.; Rawlings, M.; Gibson, D.T.; Labbe, D.; Bergeron, H.; Brousseau, R.; Lau, P.C.  
Mol. Gen. Genet. 246, 570-579, 1995  
A;Title: Identification of a membrane protein and a truncated LysR-type regulator associ  
A;Reference number: SS3993; MUID:95214617; PMID:7535376  
A;Accession: SS3994  
A;Status: preliminary  
A;Molecule type: DNA  
A;Residues: 1-100 <WAN>  
A;Cross-references: GB:U18304; NID:g747633; PIDN:AAC43317.1; PID:g747635  
A;Note: the authors translated the codon GAG for residue 71 as Asp and CGC for residue 8

Query Match 38.2%; Score 29; DB 2; Length 100;  
Best Local Similarity 50.0%; Pred. No. 4.5e+02;  
Matches 5; Conservative 4; Mismatches 1; Indels 0; Gaps 0;

Qy 6 LNLVLMANNII 15  
|||:|:|  
Db 8 LNLVLMANNAL 17

RESULT 42  
SI5650  
NADH2 dehydrogenase (ubiquinone) (EC 1.6.5.3) chain 4 - root-knot nematode (*Meloidogyne*  
C;Species: *Meloidogyne javanica*  
C;Date: 19-Mar-1997 #sequence\_revision 29-Aug-1997 #text\_change 03-Jun-2002  
C;Accession: SI5650  
R;Okamoto, R.; Chamberlin, H.M.; Macfarlane, J.L.; Wolstenholme, D.R.  
Nucleic Acids Res. 19, 1619-1626, 1991  
A;Title: Repeated sequence sets in mitochondrial DNA molecules of root knot nematodes (M  
A;Reference number: SI5649; MUID:91227158; PMID:2027769  
A;Accession: SI5650  
A;Molecule type: DNA  
A;Residues: 1-30 <OKI>  
A;Cross-references: EMBL:X57625; NID:gl3102; PID:el192325; PID:g2654262

C;Genetics:  
A;Gene: ND4  
A;Genome: mitochondrion  
A;Genetic code: SGC4  
C;Keywords: mitochondrion; NAD; oxidoreductase

Query Match 36.8%; Score 28; DB 2; Length 30;  
Best Local Similarity 41.7%; Pred. No. 1.8e+02;  
Matches 5; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

Qy 3 YDALNVLMMNII 14  
|:|:|:|:  
Db 5 YKKNVNSMYMNL 16

RESULT 43  
C90001  
hypothetical protein [imported] - Staphylococcus aureus (strain N315)  
C;Species: Staphylococcus aureus  
C;Date: 10-May-2001 #sequence\_revision 10-May-2001 #text\_change 22-Oct-2001  
C;Accession: C90001  
R;Kuroda, M.; Ohta, T.; Uchiyama, I.; Baba, T.; Yuzawa, H.; Kobayashi, I.; Cui, L.; Oguc  
ma, A.; Mizutani-Ui, Y.; Kobayashi, N.; Sawano, T.; Inoue, R.; Kaito, C.; Sekimizu, K.;  
C.; Shiba, T.; Hattori, M.; Ogasawara, N.; Hayashi, H.; Hiramatsu, K.  
Lancet 357, 1225-1240, 2001  
A;Title: Whole genome sequencing of methicillin-resistant Staphylococcus aureus.  
A;Reference number: A89759; PMID:21311952; PMID:11418146  
A;Accession: C90001  
A;Status: preliminary  
A;Molecule type: DNA  
A;Residues: 1-45 <KUR>  
A;Cross-references: GB:BA000018; PID:g13701880; PIDN:BA843172.1; GSPDB:GN00149  
A;Experimental source: strain N315  
C;Genetics:  
A;Gene: SAS068

Query Match 36.8%; Score 28; DB 2; Length 45;  
Best Local Similarity 46.2%; Pred. No. 2.9e+02;  
Matches 6; Conservative 2; Mismatches 5; Indels 0; Gaps 0;

Qy 4 DALNVLMMNIIIS 16  
|:|:|:|:|:  
Db 29 DIINTLPASNKVS 41

RESULT 44  
G69115  
hypothetical protein MTH186 - Methanobacterium thermoautotrophicum (strain Delta H)  
C;Species: Methanobacterium thermoautotrophicum  
C;Date: 05-Dec-1997 #sequence\_revision 05-Dec-1997 #text\_change 22-Oct-1999  
C;Accession: G69115  
R;Smith, D.R.; Doucette-Stamm, L.A.; Deloughery, C.; Lee, H.; Dubois, J.; Aldredge, T.;  
Qiu, D.; Spadafora, R.; Vicaire, R.; Wang, Y.; Wierzbowski, J.; Gibson, R.; Jiواني, N.  
ki, S.; Church, G.M.; Daniels, C.J.; Mao, J.; Rice, P.; Noelling, J.; Reeve, J.N.  
J. Bacteriol. 179, 7135-7155, 1997  
A;Title: Complete genome sequence of Methanobacterium thermoautotrophicum Delta H: func  
A;Reference number: A69000; PMID:98037514; PMID:9371463  
A;Accession: G69115  
A;Status: preliminary; nucleic acid sequence not shown; translation not shown  
A;Molecule type: DNA  
A;Residues: 1-60 <MTH>  
A;Cross-references: GB:AE000806; GB:AE000666; MID:g2621222; PIDN:AAB84692.1; PID:g262123  
A;Experimental source: strain Delta H  
C;Genetics:  
A;Gene: MTH186  
A;Start codon: GTG

Query Match 36.8%; Score 28; DB 2; Length 60;  
Best Local Similarity 45.5%; Pred. No. 3.9e+02;  
Matches 5; Conservative 4; Mismatches 2; Indels 0; Gaps 0;

Qy 5 ALNVLMMNII 15  
|:|:|:|:|:  
|:|:|:|:|:

Db 20 AINISIFLNII 30

RESULT 45  
AB2812  
conserved hypothetical protein Atul916 [imported] - Agrobacterium tumefaciens (strain C5  
C;Species: Agrobacterium tumefaciens  
C;Date: 11-Jan-2002 #sequence\_revision 11-Jan-2002 #text\_change 18-Nov-2002  
C;Accession: AB2812  
R;Wood, D.W.; Secubal, J.C.; Kaul, R.; Monke, D.; Chen, L.; Wood, G.E.; Chen, Y.; Woo, L.  
erage, G.; Gillet, W.; Grant, C.; Guenther, D.; Kutyavin, T.; Levy, R.; Li, M.; McClell  
; Karp, P.; Romero, P.; Zhang, S.  
Science 294, 2317-2323, 2001  
A;Authors: Yoo, H.; Tao, Y.; Biddle, P.; Jung, M.; Krespan, W.; Perry, M.; Gordon-Kamm,  
ster, E.W.  
A;Title: The Genome of the Natural Genetic Engineer Agrobacterium tumefaciens C58.  
A;Reference number: AB2577; PMID:21608550; PMID:11743193  
A;Accession: AB2812  
A;Status: preliminary  
A;Molecule type: DNA  
A;Residues: 1-63 <KUR>  
A;Cross-references: GB:AE008688; PIDN:AAL42912.1; PID:g17740367; GSPDB:GN00186  
A;Experimental source: strain C58 (Dupont)  
C;Genetics:  
A;Gene: Atul916  
A;Map position: circular chromosome

Query Match 36.8%; Score 28; DB 2; Length 63;  
Best Local Similarity 60.0%; Pred. No. 4.1e+02;  
Matches 6; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

Qy 3 YDALNVLMMAM 12  
|:|:|:|:|:  
Db 5 FDILTQLMMAM 14

Search completed: February 17, 2004, 10:58:17  
Job time : 9.66337 secs

Result No.	Score	Query Match	Length	DB	ID	Description
1	152	100.0	30	10	US-09-900-147-6	Sequence 6, Appli
2	152	100.0	37	10	US-09-900-147-1	Sequence 1, Appli
3	126	82.4	74	15	US-10-214-188-10	Sequence 10, Appl
4	101	66.4	20	10	US-09-900-147-4	Sequence 4, Appli
5	80	52.6	15	10	US-09-900-147-10	Sequence 10, Appl
6	72	47.4	15	10	US-09-900-147-3	Sequence 3, Appli
7	68	44.7	19	10	US-09-900-147-15	Sequence 15, Appl
8	67	44.1	16	10	US-09-900-147-5	Sequence 5, Appli
9	64	42.1	19	10	US-09-900-147-16	Sequence 16, Appl
10	60	39.5	19	10	US-09-900-147-17	Sequence 17, Appl
11	56.5	37.2	74	15	US-10-214-188-6	Sequence 6, Appli
12	55.5	36.5	74	15	US-10-214-188-5	Sequence 5, Appli
13	55.5	36.5	85	11	US-09-932-581-5	Sequence 5, Appli
14	55.5	36.5	85	12	US-10-338-294-5	Sequence 5, Appli
15	55.5	36.5	85	14	US-10-165-614-2	Sequence 2, Appli

```
;
; GENERAL INFORMATION:
; APPLICANT: La Thangue, Nicholas B
; APPLICANT: Bandara, Lasantha R
; TITLE OF INVENTION: Peptide antagonists of DP transcription factors
; FILE REFERENCE: 620-67
; CURRENT APPLICATION NUMBER: US/09/900,147
; CURRENT FILING DATE: 2001-07-09
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 09/308,935
; PRIOR FILING DATE: EARLIER FILING DATE: 1999-05-27
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: GB 9626589.7
; PRIOR FILING DATE: EARLIER FILING DATE: 1996-12-20
; NUMBER OF SEQ ID NOS: 18
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1
; LENGTH: 37
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Synthetic
; US-09-900-147-1

Query Match      100.0%; Score 152; DB 10; Length 37;
Best Local Similarity 100.0%; Pred. No. 9.8e-16;
Matches 30; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 YDALNVLMMNIIISKEKEIKWIGLPTNSA 30
   |||||
Db 8 YDALNVLMMNIIISKEKEIKWIGLPTNSA 37

RESULT 3
US-10-214-188-10
; Sequence 10, Application US/10214188
; Publication No. US2003002260A1
; GENERAL INFORMATION:
; APPLICANT: LA THANGUE, NICHOLAS B.
; BERNARDS, RENE
; HIJMAN, ELEANORE M.
; TITLE OF INVENTION: TRANSCRIPTION FACTOR E2F-5
; NUMBER OF SEQUENCES: 25
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: NIXON & VANDERHUYE P.C.
; STREET: 1100 NORTH GLEBE ROAD
; CITY: ARLINGTON
; STATE: VIRGINIA
; COUNTRY: U.S.A.
; ZIP: 22201-4714
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/10/214,188
; FILING DATE: 08-Aug-2002
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/894,139
; FILING DATE: 13-AUG-1997
; NAME: WILSON, MARY J.
; ATTORNEY/AGENT INFORMATION:
; REGISTRATION NUMBER: 32,955
; REFERENCE/DOCKET NUMBER: 620-22
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703) 816-4000
; TELEFAX: (703) 816-4100
; INFORMATION FOR SEQ ID NO: 10:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 74 amino acids
; TYPE: amino acid
; STRANDEDNESS: <Unknown>
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
;

; SEQUENCE DESCRIPTION: SEQ ID NO: 10:
US-10-214-188-10
Query Match      82.9%; Score 126; DB 15; Length 74;
Best Local Similarity 100.0%; Pred. No. 1.8e-11;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 YDALNVLMMNIIISKEKEIKWIGL 25
   |||||
Db 50 YDALNVLMMNIIISKEKEIKWIGL 74

RESULT 4
US-09-900-147-4
; Sequence 4, Application US/09900147
; Patent No. US20020103121A1
; GENERAL INFORMATION:
; APPLICANT: La Thangue, Nicholas B
; APPLICANT: Bandara, Lasantha R
; TITLE OF INVENTION: Peptide antagonists of DP transcription factors
; FILE REFERENCE: 620-67
; CURRENT APPLICATION NUMBER: US/09/900,147
; CURRENT FILING DATE: 2001-07-09
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 09/308,935
; PRIOR FILING DATE: EARLIER FILING DATE: 1999-05-27
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: GB 9626589.7
; PRIOR FILING DATE: EARLIER FILING DATE: 1996-12-20
; NUMBER OF SEQ ID NOS: 18
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 4
; LENGTH: 20
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Synthetic
; US-09-900-147-4

Query Match      66.4%; Score 101; DB 10; Length 20;
Best Local Similarity 100.0%; Pred. No. 2.5e-08;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 5 NVLMNMNIIISKEKEIKWIG 24
   |||||
Db 1 NVLMNMNIIISKEKEIKWIG 20

RESULT 5
US-09-900-147-10
; Sequence 10, Application US/09900147
; Patent No. US20020103121A1
; GENERAL INFORMATION:
; APPLICANT: La Thangue, Nicholas B
; APPLICANT: Bandara, Lasantha R
; TITLE OF INVENTION: Peptide antagonists of DP transcription factors
; FILE REFERENCE: 620-67
; CURRENT APPLICATION NUMBER: US/09/900,147
; CURRENT FILING DATE: 2001-07-09
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 09/308,935
; PRIOR FILING DATE: EARLIER FILING DATE: 1999-05-27
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: GB 9626589.7
; PRIOR FILING DATE: EARLIER FILING DATE: 1996-12-20
; NUMBER OF SEQ ID NOS: 18
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 10
; LENGTH: 15
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Synthetic
; US-09-900-147-10

Query Match      52.6%; Score 80; DB 10; Length 15;
Best Local Similarity 100.0%; Pred. No. 2.7e-05;
```

Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 16 EKKEIKWIGLPTNSA 30  
| | | | | | | | | | | | | | | | | |  
Db 1 EKKEIKWIGLPTNSA 15

## RESULT 6

US-09-900-147-3  
; Sequence 3, Application US/09900147  
; Patent No. US20020103121A1  
; GENERAL INFORMATION:  
; APPLICANT: La Thangue, Nicholas B  
; APPLICANT: Bandara, Lasantha R  
; TITLE OF INVENTION: Peptide antagonists of DP transcription factors  
; FILE REFERENCE: 620-67  
; CURRENT APPLICATION NUMBER: US/09/900,147  
; CURRENT FILING DATE: 2001-07-09  
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 09/308,935  
; PRIOR FILING DATE: EARLIER FILING DATE: 1999-05-27  
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: GB 9626589.7  
; PRIOR FILING DATE: EARLIER FILING DATE: 1996-12-20  
; NUMBER OF SEQ ID NOS: 18  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 3  
; LENGTH: 19  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Description of Artificial Sequence: Synthetic  
US-09-900-147-3

Query Match 47.4%; Score 72; DB 10; Length 19;  
Best Local Similarity 100.0%; Pred. No. 0.00056;  
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 YDALNVLMMNNIISK 15  
| | | | | | | | | | | | | | | | | |  
Db 5 YDALNVLMMNNIISK 19

## RESULT 7

US-09-900-147-15  
; Sequence 15, Application US/09900147  
; Patent No. US20020103121A1  
; GENERAL INFORMATION:  
; APPLICANT: La Thangue, Nicholas B  
; APPLICANT: Bandara, Lasantha R  
; TITLE OF INVENTION: Peptide antagonists of DP transcription factors  
; FILE REFERENCE: 620-67  
; CURRENT APPLICATION NUMBER: US/09/900,147  
; CURRENT FILING DATE: 2001-07-09  
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 09/308,935  
; PRIOR FILING DATE: EARLIER FILING DATE: 1999-05-27  
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: GB 9626589.7  
; PRIOR FILING DATE: EARLIER FILING DATE: 1996-12-20  
; NUMBER OF SEQ ID NOS: 18  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 15  
; LENGTH: 19  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Description of Artificial Sequence: Mutant peptide  
US-09-900-147-15

Query Match 44.7%; Score 68; DB 10; Length 19;  
Best Local Similarity 93.3%; Pred. No. 0.0022;  
Matches 14; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 YDALNVLMMNNIISK 15  
| | | | | | | | | | | | | | | | | |  
Db 5 YDALNVLMMNNIISK 19

## RESULT 8

US-09-900-147-5  
; Sequence 5, Application US/09900147  
; Patent No. US20020103121A1  
; GENERAL INFORMATION:  
; APPLICANT: La Thangue, Nicholas B  
; APPLICANT: Bandara, Lasantha R  
; TITLE OF INVENTION: Peptide antagonists of DP transcription factors  
; FILE REFERENCE: 620-67  
; CURRENT APPLICATION NUMBER: US/09/900,147  
; CURRENT FILING DATE: 2001-07-09  
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 09/308,935  
; PRIOR FILING DATE: EARLIER FILING DATE: 1999-05-27  
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: GB 9626589.7  
; PRIOR FILING DATE: EARLIER FILING DATE: 1996-12-20  
; NUMBER OF SEQ ID NOS: 18  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 5  
; LENGTH: 16  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Description of Artificial Sequence: Synthetic  
US-09-900-147-5

Query Match 44.1%; Score 67; DB 10; Length 16;  
Best Local Similarity 100.0%; Pred. No. 0.0026;  
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 YDALNVLMMNNIIS 14  
| | | | | | | | | | | | | | | | | |  
Db 3 YDALNVLMMNNIIS 16

## RESULT 9

US-09-900-147-16  
; Sequence 16, Application US/09900147  
; Patent No. US20020103121A1  
; GENERAL INFORMATION:  
; APPLICANT: La Thangue, Nicholas B  
; APPLICANT: Bandara, Lasantha R  
; TITLE OF INVENTION: Peptide antagonists of DP transcription factors  
; FILE REFERENCE: 620-67  
; CURRENT APPLICATION NUMBER: US/09/900,147  
; CURRENT FILING DATE: 2001-07-09  
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 09/308,935  
; PRIOR FILING DATE: EARLIER FILING DATE: 1999-05-27  
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: GB 9626589.7  
; PRIOR FILING DATE: EARLIER FILING DATE: 1996-12-20  
; NUMBER OF SEQ ID NOS: 18  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 16  
; LENGTH: 19  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Description of Artificial Sequence: Mutant peptide  
US-09-900-147-16

Query Match 42.1%; Score 64; DB 10; Length 19;  
Best Local Similarity 93.3%; Pred. No. 0.009;  
Matches 14; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 YDALNVLMMNNIISK 15  
| | | | | | | | | | | | | | | | | |  
Db 5 YDALNVLMMNNIISK 19

## RESULT 10

US-09-900-147-17  
; Sequence 17, Application US/09900147





```

; APPLICANT: Andrews, William H.
; APPLICANT: Foster, Christopher A.
; APPLICANT: Fraser, Stephanie
; APPLICANT: Mohammadpour, Hamid
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR MODULATING
; TITLE OF INVENTION: TELOMERASE REVERSE TRANSCRIPTASE (TEXT) EXPRESSION
; FILE REFERENCE: SIER-005
; CURRENT APPLICATION NUMBER: US/09/932,581
; CURRENT FILING DATE: 2001-08-17
; PRIOR APPLICATION NUMBER: 60/227,865
; PRIOR FILING DATE: 2000-08-24
; PRIOR APPLICATION NUMBER: 60/230,174
; PRIOR FILING DATE: 2000-09-01
; PRIOR APPLICATION NUMBER: 60/238,345
; PRIOR FILING DATE: 2000-10-05
; NUMBER OF SEQ ID NOS: 25
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 5
; LENGTH: 85
; TYPE: PRT
; ORGANISM: human
US-09-932-581-5

Query Match      36.5%; Score 55.5; DB 11; Length 85;
Best Local Similarity 44.0%; Pred. No. 0.9;
Matches 11; Conservative 6; Mismatches 7; Indels 1; Gaps

QY      ,      1 YDALNVLAMNIIISKE-KKEIKWIG 24
Db      59 YDITNVLEGIQIAKSKNHIQWLIG 83

RESULT 14
US-10-338-294-5
; Sequence 5, Application US/10338294
; Publication No. US20030171326A1
; GENERAL INFORMATION:
; APPLICANT: Andrews, William H.
; APPLICANT: Foster, Christopher A.
; APPLICANT: Fraser, Stephanie
; APPLICANT: Mohammadpour, Hamid
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR MODULATING
; TITLE OF INVENTION: TELOMERASE REVERSE TRANSCRIPTASE (TEXT) EXPRESSION
; FILE REFERENCE: SIER-005
; CURRENT APPLICATION NUMBER: US/10/338,294
; CURRENT FILING DATE: 2003-01-07
; PRIOR APPLICATION NUMBER: US/09/932,581
; PRIOR FILING DATE: 2001-08-17
; PRIOR APPLICATION NUMBER: 60/227,865
; PRIOR FILING DATE: 2000-08-24
; PRIOR APPLICATION NUMBER: 60/230,174
; PRIOR FILING DATE: 2000-09-01
; PRIOR APPLICATION NUMBER: 60/238,345
; PRIOR FILING DATE: 2000-10-05
; NUMBER OF SEQ ID NOS: 25
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 5
; LENGTH: 85
; TYPE: PRT
; ORGANISM: human
US-10-338-294-5

Query Match      36.5%; Score 55.5; DB 12; Length 85;
Best Local Similarity 44.0%; Pred. No. 0.9;
Matches 11; Conservative 6; Mismatches 7; Indels 1; Gaps

QY      1 YDALNVLAMNIIISKE-KKEIKWIG 24
Db      59 YDITNVLEGIQIAKSKNHIQWLIG 83

RESULT 15
US-10-165-614-2

```

ADRESSEE: NIXON & VANDERHYE P.C.  
STREET: 1100 NORTH GLEBE ROAD  
CITY: ARLINGTON  
STATE: VIRGINIA  
COUNTRY: U.S.A.  
ZIP: 22201-4714  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/10/214,188  
FILING DATE: 08-Aug-2002  
CLASSIFICATION: <Unknown>  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US/08/894,139  
FILING DATE: 13-AUG-1997  
ATTORNEY/AGENT INFORMATION:  
NAME: WILSON, MARY J.  
REGISTRATION NUMBER: 32,955  
REFERENCE/DOCKET NUMBER: 620-22  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (703) 816-4000  
TELEFAX: (703) 816-4100  
INFORMATION FOR SEQ ID NO: 7:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 74 amino acids  
TYPE: amino acid  
STRANDEDNESS: <Unknown>  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
SEQUENCE DESCRIPTION: SEQ ID NO: 7:  
US-10-214-188-7  
Query Match 35.2%; Score 53.5; DB 15; Length 74;  
Best Local Similarity 40.0%; Pred. No. 1.5;  
Matches 10; Conservative 7; Mismatches 7; Indels 1; Gaps 1;  
Qy 1 YDALNVLMMNIISKE-KKEIKWIG 24  
Db 49 YDITNVLEGIHLIKKSKNHVQWNG 73  
RESULT 18  
US-09-764-877-1380  
Sequence 1380, Application US/09764877  
Patent No. US20020147140A1  
GENERAL INFORMATION:  
APPLICANT: ROSEN ET AL.  
TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies  
FILE REFERENCE: PC005  
CURRENT APPLICATION NUMBER: US/09/764,877  
CURRENT FILING DATE: 2001-01-17  
Prior application data removed - refer to PALM or file wrapper  
NUMBER OF SEQ ID NOS: 4031  
SOFTWARE: PatentIn Ver. 2.0  
SEQ ID NO 1380  
LENGTH: 29  
TYPE: PRT  
ORGANISM: Homo sapiens  
US-09-764-877-1380  
Query Match 34.2%; Score 52; DB 10; Length 29;  
Best Local Similarity 62.5%; Pred. No. 0.93;  
Matches 10; Conservative 2; Mismatches 4; Indels 0; Gaps 0;  
Qy 10 MNIIISKEKKEIKWIGL 25  
Db 13 MKIFSKEKKKGWPGM 28  
RESULT 19  
US-09-764-877-1380  
Sequence 1380, Application US/10242515  
Patent No. US20040009488A1  
GENERAL INFORMATION:  
APPLICANT: ROSEN ET AL.  
TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies  
FILE REFERENCE: PC005C1  
CURRENT APPLICATION NUMBER: US/10/242,515  
CURRENT FILING DATE: 2002-09-13  
PRIOR FILING DATE: 2001-01-17  
PRIOR APPLICATION NUMBER: 60/179,065  
PRIOR FILING DATE: 2000-01-31  
PRIOR APPLICATION NUMBER: 60/180,628  
PRIOR FILING DATE: 2000-02-04  
PRIOR APPLICATION NUMBER: 60/214,886  
PRIOR FILING DATE: 2000-06-28  
PRIOR APPLICATION NUMBER: 60/217,487  
PRIOR FILING DATE: 2000-07-11  
PRIOR APPLICATION NUMBER: 60/225,758  
PRIOR FILING DATE: 2000-08-14  
PRIOR APPLICATION NUMBER: 60/220,963  
PRIOR FILING DATE: 2000-07-26  
PRIOR APPLICATION NUMBER: 60/217,496  
PRIOR FILING DATE: 2000-07-11  
PRIOR APPLICATION NUMBER: 60/225,447  
PRIOR FILING DATE: 2000-08-14  
PRIOR APPLICATION NUMBER: 60/218,290  
PRIOR FILING DATE: 2000-07-14  
Remaining Prior Application data removed - See File Wrapper or PALM.  
NUMBER OF SEQ ID NOS: 4031  
SOFTWARE: PatentIn Ver. 2.0  
SEQ ID NO 1380  
LENGTH: 29  
TYPE: PRT  
ORGANISM: Homo sapiens  
US-10-242-515-1380  
Query Match 34.2%; Score 52; DB 12; Length 29;  
Best Local Similarity 62.5%; Pred. No. 0.93;  
Matches 10; Conservative 2; Mismatches 4; Indels 0; Gaps 0;  
Qy 10 MNIIISKEKKEIKWIGL 25  
Db 13 MKIFSKEKKKGWPGM 28  
RESULT 20  
US-10-214-188-9  
Sequence 9, Application US/10214188  
Patent No. US2003002260A1  
GENERAL INFORMATION:  
APPLICANT: LA THANGUE, NICHOLAS B.  
BERNARDS, RENE  
HUMANS, ELEANORE M.  
TITLE OF INVENTION: TRANSCRIPTION FACTOR E2F-5  
NUMBER OF SEQUENCES: 25  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: NIXON & VANDERHYE P.C.  
STREET: 1100 NORTH GLEBE ROAD  
CITY: ARLINGTON  
STATE: VIRGINIA  
COUNTRY: U.S.A.  
ZIP: 22201-4714  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/10/214,188  
FILING DATE: 08-Aug-2002  
CLASSIFICATION: <Unknown>

; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US/08/894,139  
; FILING DATE: 13-AUG-1997  
; ATTORNEY/AGENT INFORMATION:  
; NAME: WILSON, MARY J.  
; REGISTRATION NUMBER: 32,955  
; REFERENCE/DOCKET NUMBER: 620-22  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (703) 816-4000  
; TELEFAX: (703) 816-4100  
; INFORMATION FOR SEQ ID NO: 9:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 75 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: <Unknown>  
; TOPOLOGY: linear  
; MOLECULE TYPE: peptide  
; SEQUENCE DESCRIPTION: SEQ ID NO: 9:  
US-10-214-188-9

Query Match 33.9%; Score 51.5; DB 15; Length 75;  
Best Local Similarity 42.3%; Pred. No. 3.1;  
Matches 11; Conservative 6; Mismatches 8; Indels 1; Gaps 1;

Qy 1 YDALNVLMMNNIISKE-KKEIKWIGL 25  
Db 50 YDITNVLEGIGLIEKKSKNSIQWKG 75

## RESULT 21

US-09-932-581-6  
; Sequence 6, Application US/09932581  
; Publication No. US20030050264A1  
; GENERAL INFORMATION:

; APPLICANT: Andrews, William H.  
; APPLICANT: Foster, Christopher A.  
; APPLICANT: Fraser, Stephanie  
; APPLICANT: Mohammadpour, Hamid  
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR MODULATING  
; TITLE OF INVENTION: TELOMERASE REVERSE TRANSCRIPTASE (TERT) EXPRESSION  
; FILE REFERENCE: SIER-005  
; CURRENT APPLICATION NUMBER: US/09/932,581  
; CURRENT FILING DATE: 2001-08-17  
; PRIOR APPLICATION NUMBER: 60/227,865  
; PRIOR FILING DATE: 2000-08-24  
; PRIOR APPLICATION NUMBER: 60/230,174  
; PRIOR FILING DATE: 2000-09-01  
; PRIOR APPLICATION NUMBER: 60/238,345  
; PRIOR FILING DATE: 2000-10-05  
; NUMBER OF SEQ ID NOS: 25  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 6  
; LENGTH: 76  
; TYPE: PRT  
; ORGANISM: human  
US-09-932-581-6

Query Match 33.2%; Score 50.5; DB 11; Length 76;  
Best Local Similarity 42.3%; Pred. No. 4.5;  
Matches 11; Conservative 5; Mismatches 9; Indels 1; Gaps 1;

Qy 1 YDALNVLMMNNIISKE-KKEIKWIGL 25  
Db 49 YDITNVLEGIGLIEKKSKNSIQWKG 74

## RESULT 22

US-10-338-294-6  
; Sequence 6, Application US/10338294  
; Publication No. US20030171326A1  
; GENERAL INFORMATION:  
; APPLICANT: Andrews, William H.  
; APPLICANT: Foster, Christopher A.

; APPLICANT: Fraser, Stephanie  
; APPLICANT: Mohammadpour, Hamid  
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR MODULATING  
; TITLE OF INVENTION: TELOMERASE REVERSE TRANSCRIPTASE (TERT) EXPRESSION  
; FILE REFERENCE: SIER-005  
; CURRENT APPLICATION NUMBER: US/10/338,294  
; CURRENT FILING DATE: 2003-01-07  
; PRIOR APPLICATION NUMBER: US/09/932,581  
; PRIOR FILING DATE: 2001-08-17  
; PRIOR APPLICATION NUMBER: 60/227,865  
; PRIOR FILING DATE: 2000-08-24  
; PRIOR APPLICATION NUMBER: 60/230,174  
; PRIOR FILING DATE: 2000-09-01  
; PRIOR APPLICATION NUMBER: 60/238,345  
; PRIOR FILING DATE: 2000-10-05  
; NUMBER OF SEQ ID NOS: 25  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 6  
; LENGTH: 76  
; TYPE: PRT  
; ORGANISM: human  
US-10-338-294-6

Query Match 33.2%; Score 50.5; DB 12; Length 76;  
Best Local Similarity 42.3%; Pred. No. 4.5;  
Matches 11; Conservative 5; Mismatches 9; Indels 1; Gaps 1;

Qy 1 YDALNVLMMNNIISKE-KKEIKWIGL 25  
Db 49 YDITNVLEGIGLIEKKSKNSIQWKG 74

## RESULT 23

US-10-165-614-3  
; Sequence 3, Application US/10165614  
; Publication No. US20020193289A1  
; GENERAL INFORMATION:  
; APPLICANT: Andrews, William H.  
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR MODULATING  
; TITLE OF INVENTION: TELOMERASE REVERSE TRANSCRIPTASE (TERT) EXPRESSION  
; FILE REFERENCE: SIER-018  
; CURRENT APPLICATION NUMBER: US/10/165,614  
; CURRENT FILING DATE: 2002-06-25  
; PRIOR APPLICATION NUMBER: 60/296,992  
; PRIOR FILING DATE: 2001-06-07  
; NUMBER OF SEQ ID NOS: 4  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 3  
; LENGTH: 76  
; TYPE: PRT  
; ORGANISM: homo sapiens  
US-10-165-614-3

Query Match 33.2%; Score 50.5; DB 14; Length 76;  
Best Local Similarity 42.3%; Pred. No. 4.5;  
Matches 11; Conservative 5; Mismatches 9; Indels 1; Gaps 1;

Qy 1 YDALNVLMMNNIISKE-KKEIKWIGL 25  
Db 49 YDITNVLEGIGLIEKKSKNSIQWKG 74

## RESULT 24

US-10-029-386-29071  
; Sequence 29071, Application US/10029386  
; Publication No. US20030194704A1  
; GENERAL INFORMATION:  
; APPLICANT: Penn, Sharron G.  
; APPLICANT: Rank, David R.  
; APPLICANT: Hanzel, David K.  
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR G  
; TITLE OF INVENTION: EXPRESSION ANALYSIS TWO  
; FILE REFERENCE: AEOMICA-X-2

; CURRENT APPLICATION NUMBER: US/10/029,386  
; CURRENT FILING DATE: 2001-12-20  
; NUMBER OF SEQ ID NOS: 34288  
; SOFTWARE: Anomax Sequence Listing Engine vers. 1.1  
; SEQ ID NO 29071  
; LENGTH: 76  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
; FEATURE:  
; OTHER INFORMATION: MAP TO CHR20.1  
; OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 6.1  
; OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 5.4  
; OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 4.6  
; OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 3.3  
; OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 4.8  
; OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 4.4  
; OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 4  
; OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 5.4  
; OTHER INFORMATION: EXPRESSED IN SWISSPROT HIT: Q01094, EVALUE 1.00e-36  
US-10-029-386-29071

Query Match 32.6%; Score 49.5; DB 12; Length 76;  
Best Local Similarity 41.7%; Pred. No. 6.4;  
Matches 10; Conservative 6; Mismatches 7; Indels 1; Gaps 1;

Qy 1 YDALNVLMMNIISKE-KKEIKWI 23  
||| : : : : :  
Db 53 YDITNVLEGIQIAKSKNHIQWL 76

## RESULT 25

US-09-864-761-45697  
; Sequence 45697, Application US/09864761  
; Patent No. US20020048763A1  
; GENERAL INFORMATION:  
; APPLICANT: Penn, Sharron G.  
; APPLICANT: Rank, David R.  
; APPLICANT: Hanzel, David K.  
; APPLICANT: Chen, Wensheng  
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR  
; FILE REFERENCE: Acmica-X-1  
; CURRENT APPLICATION NUMBER: US/09/864,761  
; CURRENT FILING DATE: 2001-05-23  
; PRIOR APPLICATION NUMBER: US 60/180,312  
; PRIOR FILING DATE: 2000-02-04  
; PRIOR APPLICATION NUMBER: US 60/207,456  
; PRIOR FILING DATE: 2000-05-26  
; PRIOR APPLICATION NUMBER: US 09/632,366  
; PRIOR FILING DATE: 2000-08-03  
; PRIOR APPLICATION NUMBER: GB 24263.6  
; PRIOR FILING DATE: 2000-10-04  
; PRIOR APPLICATION NUMBER: US 60/236,359  
; PRIOR FILING DATE: 2000-09-27  
; PRIOR APPLICATION NUMBER: PCT/US01/00666  
; PRIOR FILING DATE: 2001-01-30  
; PRIOR APPLICATION NUMBER: PCT/US01/00667  
; PRIOR FILING DATE: 2001-01-30  
; PRIOR APPLICATION NUMBER: PCT/US01/00664  
; PRIOR FILING DATE: 2001-01-30  
; PRIOR APPLICATION NUMBER: PCT/US01/00669  
; PRIOR FILING DATE: 2001-01-30  
; PRIOR APPLICATION NUMBER: PCT/US01/00665  
; PRIOR FILING DATE: 2001-01-30  
; PRIOR APPLICATION NUMBER: PCT/US01/00668  
; PRIOR FILING DATE: 2001-01-30  
; PRIOR APPLICATION NUMBER: PCT/US01/00663  
; PRIOR FILING DATE: 2001-01-30  
; PRIOR APPLICATION NUMBER: PCT/US01/00662  
; PRIOR FILING DATE: 2001-01-30  
; PRIOR APPLICATION NUMBER: PCT/US01/00661  
; PRIOR FILING DATE: 2001-01-30  
; PRIOR APPLICATION NUMBER: PCT/US01/00670

; PRIOR FILING DATE: 2001-01-30  
; PRIOR APPLICATION NUMBER: US 60/234,687  
; PRIOR FILING DATE: 2000-09-21  
; PRIOR APPLICATION NUMBER: US 09/608,408  
; PRIOR FILING DATE: 2000-06-30  
; PRIOR APPLICATION NUMBER: US 09/774,203  
; PRIOR FILING DATE: 2001-01-29  
; NUMBER OF SEQ ID NOS: 49117  
; SOFTWARE: Anomax Sequence Listing Engine vers. 1.1  
; SEQ ID NO 45697  
; LENGTH: 96  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
; FEATURE:  
; OTHER INFORMATION: MAP TO AC021804.3  
; OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 0.89  
; OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 0.99  
; OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 0.9  
; OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 0.92  
; OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 0.88  
; OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 0.75  
; OTHER INFORMATION: EST HUMAN HIT: B8880658.1, EVALUE 2.00e-51  
; OTHER INFORMATION: SWISSPROT HIT: Q61501, EVALUE 5.00e-06  
US-09-864-761-45697

Query Match 30.6%; Score 46.5; DB 9; Length 96;  
Best Local Similarity 36.0%; Pred. No. 23;  
Matches 9; Conservative 8; Mismatches 7; Indels 1; Gaps 1;

Qy 1 YDALNVLMMNIISKE-EKKEIKWIG 24  
||| : : : : :  
Db 7 YDIVNVLESILHLSRVAKNQYGMHG 31

## RESULT 26

US-10-214-188-8  
; Sequence 8, Application US/10214188  
; Publication No. US2003002260A1  
; GENERAL INFORMATION:  
; APPLICANT: LA THANGUE, NICHOLAS B.  
; BERNARDS, RENE  
; HIGMANS, ELEANORE M.  
; TITLE OF INVENTION: TRANSCRIPTION FACTOR E2F-5  
; NUMBER OF SEQUENCES: 25  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: NIXON & VANDERHVE P.C.  
; STREET: 1100 NORTH GLEBE ROAD  
; CITY: ARLINGTON  
; STATE: VIRGINIA  
; COUNTRY: U.S.A.  
; ZIP: 22201-4714  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/10/214,188  
; FILING DATE: 08-Aug-2002  
; CLASSIFICATION: <Unknown>  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US/08/894,139  
; FILING DATE: 13-AUG-1997  
; ATTORNEY/AGENT INFORMATION:  
; NAME: WILSON, MARY J.  
; REGISTRATION NUMBER: 32,955  
; REFERENCE/DOCKET NUMBER: 620-22  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (703) 816-4000  
; TELEFAX: (703) 816-4100  
; INFORMATION FOR SEQ ID NO: 8:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 69 amino acids

```
; TYPE: amino acid
; STRANDEDNESS: <Unknown>
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; SEQUENCE DESCRIPTION: SEQ ID NO: 8:
US-10-214-188-8

Query Match      29.9%; Score 45.5; DB 15; Length 69;
Best Local Similarity 38.5%; Pred. No. 23;
Matches 10; Conservative 5; Mismatches 10; Indels 1; Gaps 1;

QY 1 YDALNVLMMNIIISKE-KKEIKWIGL 25
Db 44 YDITNVLEGIGLIEKSKSNSTQMRGV 69

RESULT 27
US-09-900-147-2
; Sequence 2, Application US/09900147
; Patent No. US20020103121A1
; GENERAL INFORMATION:
; APPLICANT: La Thangue, Nicholas B
; APPLICANT: Bandara, Laseantha R
; TITLE OF INVENTION: Peptide antagonists of DP transcription factors
; FILE REFERENCE: 620-67
; CURRENT APPLICATION NUMBER: US/09/900,147
; PRIOR FILING DATE: 2001-07-09
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 09/308,935
; PRIOR FILING DATE: EARLIER FILING DATE: 1999-05-27
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: GB 9626589.7
; PRIOR FILING DATE: EARLIER FILING DATE: 1996-12-20
; NUMBER OF SEQ ID NOS: 18
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2
; LENGTH: 9
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Synthetic
US-09-900-147-2

Query Match      27.6%; Score 42; DB 10; Length 9;
Best Local Similarity 100.0%; Pred. No. 7e+05;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 5 NVLMAMNII 13
Db 1 NVLMAMNII 9

RESULT 28
US-10-097-065-532
; Sequence 532, Application US/10097065
; Publication No. US20030055236A1
; GENERAL INFORMATION:
; APPLICANT: Moore, Paul A. et al.
; TITLE OF INVENTION: 110 Human Secreted Proteins
; FILE REFERENCE: P2021p1
; CURRENT APPLICATION NUMBER: US/10/097,065
; CURRENT FILING DATE: 2002-03-14
; PRIOR APPLICATION NUMBER: PCT/US98/27059
; PRIOR FILING DATE: 1998-12-17
; PRIOR APPLICATION NUMBER: 60/070,923
; PRIOR FILING DATE: 1997-12-18
; PRIOR APPLICATION NUMBER: 60/068,007
; PRIOR FILING DATE: 1997-12-18
; PRIOR APPLICATION NUMBER: 60/068,057
; PRIOR FILING DATE: 1997-12-18
; PRIOR APPLICATION NUMBER: 60/068,006
; PRIOR FILING DATE: 1997-12-18
; PRIOR APPLICATION NUMBER: 60/068,369
; PRIOR FILING DATE: 1997-12-19
; PRIOR APPLICATION NUMBER: 60/068,367

; PRIOR FILING DATE: 1997-12-19
; PRIOR APPLICATION NUMBER: 60/068,368
; PRIOR FILING DATE: 1997-12-19
; PRIOR APPLICATION NUMBER: 60/068,169
; PRIOR FILING DATE: 1997-12-19
; PRIOR APPLICATION NUMBER: 60/068,053
; PRIOR FILING DATE: 1997-12-18
; PRIOR APPLICATION NUMBER: 60/068,064
; PRIOR FILING DATE: 1997-12-18
; PRIOR APPLICATION NUMBER: 60/068,054
; PRIOR FILING DATE: 1997-12-18
; PRIOR APPLICATION NUMBER: 60/068,008
; PRIOR FILING DATE: 1997-12-18
; PRIOR APPLICATION NUMBER: 60/068,365
; PRIOR FILING DATE: 1997-12-19
; NUMBER OF SEQ ID NOS: 672
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 532
; LENGTH: 57
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-097-065-532

Query Match      27.6%; Score 42; DB 15; Length 57;
Best Local Similarity 29.2%; Pred. No. 63;
Matches 7; Conservative 7; Mismatches 10; Indels 0; Gaps 0;

QY 5 NVLMAMNIIISKEKKEIKWIGLPTN 28
Db 1 NVCVSKNTFRVRPIKWKIPAN 24

RESULT 29
US-09-263-811-9
; Sequence 9, Application US/09263811
; Patent No. US20020058610A1
; GENERAL INFORMATION:
; APPLICANT: Jian Ni et al.
; TITLE OF INVENTION: Mammary Transforming Protein
; NUMBER OF SEQUENCES: 9
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: HUMAN GENOME SCIENCES, INC
; STREET: 9410 KEY WEST AVENUE
; CITY: ROCKVILLE
; STATE: MARYLAND
; COUNTRY: USA
; ZIP: 20850
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5 INCH DISKETTE
; COMPUTER: IBM PS/2
; OPERATING SYSTEM: MS-DOS
; SOFTWARE: WORD PERFECT 5.1
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/263,811
; FILING DATE: MAR-08-1999
; CLASSIFICATION: 536
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/743,975
; FILING DATE: 01-NOV-1996
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 60/006,187
; FILING DATE: 02 NOVEMBER 1995
; ATTORNEY/AGENT INFORMATION:
; NAME: MICHELE M. WALES
; REGISTRATION NUMBER: 43,975
; REFERENCE/DOCKET NUMBER: PF212D1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 301-309-8504
; TELEFAX: 301-309-8439
; INFORMATION FOR SEQ ID NO: 9:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 61 AMINO ACIDS
; TYPE: AMINO ACID
```

```
; STRANDEDNESS:
; TOPOLOGY: LINEAR
; MOLECULE TYPE: PROTEIN
US-09-263-811-9

Query Match      27.6%; Score 42; DB 9; Length 61;
Best Local Similarity 33.3%; Pred. No. 68;
Matches 7; Conservative 9; Mismatches 3; Indels 3; Gaps 1;

QY      6 VLMAMNIISKEKEIKWIGLP 26
      |: :|:|:|:|:|:|:|:|:|:|
Db      14 VVFSINLLSRPERE--WEGMP 32

RESULT 30
US-10-259-520-9
; Sequence 9, Application US/10259520
; Publication No. US20030027989A1
; GENERAL INFORMATION:
; APPLICANT: Jian Ni et al.
; TITLE OF INVENTION: Mammary Transforming Protein
; NUMBER OF SEQUENCES: 9
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: HUMAN GENOME SCIENCES, INC
; STREET: 9410 KEY WEST AVENUE
; CITY: ROCKVILLE
; STATE: MARYLAND
; COUNTRY: USA
; ZIP: 20850
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5 INCH DISKETTE
; COMPUTER: IBM PS/2
; OPERATING SYSTEM: MS-DOS
; SOFTWARE: WORD PERFECT 5.1
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/10/259,520
; FILING DATE: 30-Sep-2002
; CLASSIFICATION: <unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/09/263,811
; FILING DATE: MAR-08-1999
; APPLICATION NUMBER: 08/743,975
; FILING DATE: 01-NOV-1996
; APPLICATION NUMBER: 60/006,187
; FILING DATE: 02 NOVEMBER 1995
; ATTORNEY/AGENT INFORMATION:
; NAME: MICHELE M. WALES
; REGISTRATION NUMBER: 43,975
; REFERENCE/DOCKET NUMBER: PF212D1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 301-309-8504
; TELEFAX: 301-309-8439
; INFORMATION FOR SEQ ID NO: 9:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 61 AMINO ACIDS
; TYPE: AMINO ACID
; STRANDEDNESS: <Unknown>
; TOPOLOGY: LINEAR
; MOLECULE TYPE: PROTEIN
; SEQUENCE DESCRIPTION: SEQ ID NO: 9:
US-10-259-520-9

Query Match      27.6%; Score 42; DB 15; Length 61;
Best Local Similarity 33.3%; Pred. No. 68;
Matches 7; Conservative 9; Mismatches 3; Indels 3; Gaps 1;

QY      6 VLMAMNIISKEKEIKWIGLP 26
      |: :|:|:|:|:|:|:|:|:|:|
Db      14 VVFSINLLSRPERE--WEGMP 32

RESULT 31
US-10-083-357-694
```

RESULT 33

US-09-864-761-34417

Sequence 34417, Application US/09864761

Patent No. US20020048763A1

GENERAL INFORMATION:

APPLICANT: Penn, Sharon G.

APPLICANT: Rank, David R.

APPLICANT: Hanzel, David K.

APPLICANT: Chen, Wensheng

TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR

TITLE OF INVENTION: GENE EXPRESSION ANALYSIS BY MICROARRAY

FILE REFERENCE: Aemica-x-1

CURRENT APPLICATION NUMBER: US/09/864,761

CURRENT FILING DATE: 2001-05-23

PRIOR APPLICATION NUMBER: US 60/180,312

PRIOR FILING DATE: 2000-02-04

PRIOR APPLICATION NUMBER: US 60/207,456

PRIOR FILING DATE: 2000-05-26

PRIOR APPLICATION NUMBER: US 09/632,366

PRIOR FILING DATE: 2000-08-03

PRIOR APPLICATION NUMBER: GB 24263.6

PRIOR FILING DATE: 2000-10-04

PRIOR APPLICATION NUMBER: US 60/236,359

PRIOR FILING DATE: 2000-09-27

PRIOR APPLICATION NUMBER: PCT/US01/00666

PRIOR FILING DATE: 2001-01-30

PRIOR APPLICATION NUMBER: PCT/US01/00667

PRIOR FILING DATE: 2001-01-30

PRIOR APPLICATION NUMBER: PCT/US01/00664

PRIOR FILING DATE: 2001-01-30

PRIOR APPLICATION NUMBER: PCT/US01/00669

PRIOR FILING DATE: 2001-01-30

PRIOR APPLICATION NUMBER: PCT/US01/00665

PRIOR FILING DATE: 2001-01-30

PRIOR APPLICATION NUMBER: PCT/US01/00668

PRIOR FILING DATE: 2001-01-30

PRIOR APPLICATION NUMBER: PCT/US01/00663

PRIOR FILING DATE: 2001-01-30

PRIOR APPLICATION NUMBER: PCT/US01/00662

PRIOR FILING DATE: 2001-01-30

PRIOR APPLICATION NUMBER: PCT/US01/00661

```
Query Match      27.0%; Score 41; DB 12; Length 63;
Best Local Similarity 46.2%; Pred. No. 1e-02;
Matches 6; Conservative 4; Mismatches 3; Indels 0; Gaps 0;
```

```

: APPLICANT: Duwick, Jonathan P.
: APPLICANT: Navarro Acevedo, Pedro A.
: APPLICANT: Simmons, Carl R.
: TITLE OF INVENTION: Maize Basal Layer Antimicrobial Protein Polynucleotides and Methods
: FILE REFERENCE: 1406
: CURRENT APPLICATION NUMBER: US/10/205,072
: CURRENT FILING DATE: 2002-07-24
: PRIOR APPLICATION NUMBER: 60/308,045
: PRIOR FILING DATE: 2001-07-26
: NUMBER OF SEQ ID NOS: 30
: SOFTWARE: FastSeq for Windows Version 4.0
: SEQ ID NO 16
: LENGTH: 93
: TYPE: PRT
: ORGANISM: Zea mays
: US-10-205-072-16

Query Match          26.6%; Score 40.5; DB 15; Length 93;
Best Local Similarity 44.4%; Pred. No. 1.8e+02;
Matches 12; Conservative 4; Mismatches 6; Indels 5; Gaps 1;

QY      6 VLMA-----MNIISKEKEIKWIGLPT 27
      |||  :|||  |  :|  :|||
DB      19 VLLASCVTHAIIISGETEEVSNIGSPT 45

RESULT 37
US-09-820-408-4
: Sequence 4, Application US/09820408
: Patent No. US20020102701A1
: GENERAL INFORMATION:
: APPLICANT: Lonetto, Michael A.
: TITLE OF INVENTION: l1cd1
: NUMBER OF SEQUENCES: 4
: CORRESPONDENCE ADDRESS:
: ADDRESSER: Dechert, Price & Rhoads

```



```
; NAME: Falk, Stephen T
; REGISTRATION NUMBER: 36,795
; REFERENCE/DOCKET NUMBER: GM50019
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 215-994-2488
; TELEFAX: 215-994-2222
; TELEX: <Unknown>
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 100 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; SEQUENCE DESCRIPTION: SEQ ID NO: 4:
US-09-820-473--4

Query Match      26.3%; Score 40; DB 10; Length 100;
Best Local Similarity 42.1%; Pred. No. 2.3e+02;
Matches 8; Conservative 4; Mismatches 7; Indels 0; Gaps 0;

QY 10 MNIISKEKEIKWIGLPTN 28
Db 67 INIFQEKSKYKLLSLETD 85

RESULT 39
US-10-029-386-28032
; Sequence 28032, Application US/10029386
; Publication No. US20030194704A1
; GENERAL INFORMATION:
; APPLICANT: Penn, Sharron G.
; APPLICANT: Rank, David R.
; APPLICANT: Hanzel, David K.
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR G
; TITLE OF INVENTION: EXPRESSION ANALYSIS TWO
; FILE REFERENCE: AEOMICA-X-2
; CURRENT APPLICATION NUMBER: US/10/029,386
; NUMBER OF SEQ ID NOS: 34288
; SOFTWARE: Annonmax Sequence Listing Engine vers. 1.1
; SEQ ID NO 28032
; LENGTH: 64
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: MAP TO AL159176.2
; OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 0.54
; OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 0.68
; OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 0.54
; OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 0.5
US-10-029-386-28032

Query Match      25.7%; Score 39; DB 12; Length 64;
Best Local Similarity 35.3%; Pred. No. 2e+02;
Matches 12; Conservative 7; Mismatches 7; Indels 8; Gaps 2;

QY 4 LNVLMAM-----NIISKEKEIKWI-GLPTNS 29
Db 18 LDVLMVLSIVFCSLMFYKAEIWFLSGRNS 51

RESULT 40
US-09-978-303-16
; Sequence 16, Application US/09978303
; Publication No. US20030049728A1
; GENERAL INFORMATION:
; APPLICANT: Julius, David J.
; APPLICANT: Caterina, Michael J.
; APPLICANT: Brake, Anthony J.
; TITLE OF INVENTION: Nucleic acid sequences encoding
; TITLE OF INVENTION: capsaicin receptor and capsaicin receptor-related
; TITLE OF INVENTION: polypeptides and uses thereof
; FILE REFERENCE: UCAL084CON
```

```
; CURRENT APPLICATION NUMBER: US/09/978,303
; CURRENT FILING DATE: 2001-10-15
; PRIOR APPLICATION NUMBER: 09/235,451
; PRIOR FILING DATE: 1999-01-22
; PRIOR APPLICATION NUMBER: 60/072,151
; PRIOR FILING DATE: 1998-01-22
; PRIOR APPLICATION NUMBER: 08/915,461
; PRIOR FILING DATE: 1997-08-20
; NUMBER OF SEQ ID NOS: 48
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 16
; LENGTH: 75
; TYPE: PRT
; ORGANISM: Caliphora
US-09-978-303-16
```

```
Query Match      25.7%; Score 39; DB 11; Length 75;
Best Local Similarity 18.8%; Pred. No. 2.4e+02;
Matches 6; Conservative 10; Mismatches 6; Indels 10; Gaps 1;
```

```
QY 1 YDALNVLMMNI-----ISKEKKEIKW 22
Db 35 YSVNIIVLLNMLIAMMSNSYQIISERADVEW 66
```

```
RESULT 41
US-09-978-303-17
; Sequence 17, Application US/09978303
; Publication No. US20030049728A1
; GENERAL INFORMATION:
; APPLICANT: Julius, David J.
; APPLICANT: Caterina, Michael J.
; APPLICANT: Brake, Anthony J.
; TITLE OF INVENTION: Nucleic acid sequences encoding
; TITLE OF INVENTION: capsaicin receptor and capsaicin receptor-related
; TITLE OF INVENTION: polypeptides and uses thereof
; FILE REFERENCE: UCAL084CON
; CURRENT APPLICATION NUMBER: US/09/978,303
; CURRENT FILING DATE: 2001-10-15
; PRIOR APPLICATION NUMBER: 09/235,451
; PRIOR FILING DATE: 1999-01-22
; PRIOR APPLICATION NUMBER: 60/072,151
; PRIOR FILING DATE: 1998-01-22
; PRIOR APPLICATION NUMBER: 08/915,461
; PRIOR FILING DATE: 1997-08-20
; NUMBER OF SEQ ID NOS: 48
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 17
; LENGTH: 75
; TYPE: PRT
; ORGANISM: D. melanogaster
US-09-978-303-17
```

```
Query Match      25.7%; Score 39; DB 11; Length 75;
Best Local Similarity 18.8%; Pred. No. 2.4e+02;
Matches 6; Conservative 10; Mismatches 6; Indels 10; Gaps 1;
```

```
QY 1 YDALNVLMMNI-----ISKEKKEIKW 22
Db 35 YSVNIIVLLNMLIAMMSNSYQIISERADVEW 66
```

```
RESULT 42
US-10-407-920-19
; Sequence 19, Application US/10407920
; Publication No. US20040006797A1
; GENERAL INFORMATION:
; APPLICANT: Shi, Lifang
; TITLE OF INVENTION: MYB Transcription Factors and Uses for Crop Improvement
; FILE REFERENCE: 38-21(52703)A
; CURRENT APPLICATION NUMBER: US/10/407,920
; CURRENT FILING DATE: 2003-06-06
; PRIOR APPLICATION NUMBER: US/60/370,759
```

```

; PRIOR FILING DATE: 2000-10-04
; PRIOR APPLICATION NUMBER: US 60/236,359
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: PCT/US01/00666
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00667
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00664
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00669
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00665
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00668
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00663
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00662
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00661
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00670
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: US 60/234,687
; PRIOR FILING DATE: 2000-09-21
; PRIOR APPLICATION NUMBER: US 09/608,408
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: US 09/774,203
; PRIOR FILING DATE: 2001-01-29
; NUMBER OF SEQ ID NOS: 49117
; SOFTWARE: Annomax Sequence Listing Engine vers. 1.1
; SEQ ID NO 43984
; LENGTH: 31
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: MAP TO AC005386.1
; OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 0.68
; OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 0.79
; OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 0.86
; OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 0.76
; OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 0.74
; OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 0.72
; US-09-864-761-43984

Query Match 25.0%; Score 38; DB 9; Length 31;
Best Local Similarity 50.0%; Pred. No. 1.3e+02;
Matches 6; Conservative 4; Mismatches 2; Indels

Qy 11 NIISKEKEIKW 22
|:|:|:|:|
Db 7 NMISKQSRKKW 18

RESULT 45
US-10-351-641-1259
; Sequence 1259, Application US/10351641
; Publication No. US20030186874A1
; GENERAL INFORMATION:
; APPLICANT: Barney, S.
; APPLICANT: Guthrie, K.
; APPLICANT: Merutka, G.
; APPLICANT: Anwer, M.
; APPLICANT: Lambert, D.
; TITLE OF INVENTION: HYBRID POLYPEPTIDES WITH ENHANCED PHARM
; TITLE OF INVENTION: PROPERTIES
; FILE REFERENCE: 7872-100
; CURRENT APPLICATION NUMBER: US/10/351,641
; CURRENT FILING DATE: 2003-01-24
; PRIOR APPLICATION NUMBER: 09/350,641
; PRIOR FILING DATE: 1999-07-09
; PRIOR APPLICATION NUMBER: 09/315,304
; PRIOR FILING DATE: 1999-05-20

```

```
; PRIOR APPLICATION NUMBER: 09/082,279
; PRIOR FILING DATE: 1998-05-20
; NUMBER OF SEQ ID NOS: 1757
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 1259
; LENGTH: 39
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Core polypeptide
US-10-351-641-1259
```

```
Query Match      25.0%; Score 38; DB 12; Length 39;
Best Local Similarity 35.7%; Pred. No. 1.7e+02;
Matches 10; Conservative 3; Mismatches 9; Indels 6; Gaps 1;
```

```
QY      4 LNVLMAMNIISKEKE-----IKWIGL 25
      :  :  :  :  :  :  :  :  :  :  :  :
Db      8 ITALLEQAQIQEKNEYELQKLIKWAGL 35
```

```
Search completed: February 17, 2004, 11:02:22
Job time : 44.9604 secs
```

THIS PAGE BLANK (JEP70)

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: February 17, 2004, 10:50:13 ; Search time 21.6832 Seconds  
(without alignments)  
58.540 Million cell updates/sec

Title: US-09-900-147-6  
Perfect score: 152  
Sequence: 1 YDALNVLMMANNIISKKEIKWIGLPTNSA 30

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 328717 seqs, 42310858 residues

Total number of hits satisfying chosen parameters: 228043

Minimum DB seq length: 0  
Maximum DB seq length: 100

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Issued Patents AA:\*

- 1: /cgn2\_6/ptodata/1/iaa/5A COMB.pep.\*
- 2: /cgn2\_6/ptodata/1/iaa/5B COMB.pep.\*
- 3: /cgn2\_6/ptodata/1/iaa/6A COMB.pep.\*
- 4: /cgn2\_6/ptodata/1/iaa/6B COMB.pep.\*
- 5: /cgn2\_6/ptodata/1/iaa/PTUS COMB.pep.\*
- 6: /cgn2\_6/ptodata/1/iaa/backfiles1.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	152	100.0	30	3	US-09-308-935-6
2	152	100.0	37	2	US-09-308-935-1
3	152	100.0	72	2	US-08-428-131-11
4	152	100.0	72	3	US-09-078-596-11
5	126	82.9	74	4	US-08-894-139-10
6	101	66.4	20	3	US-09-308-935-4
7	80	52.6	15	3	US-09-308-935-10
8	72	47.4	19	3	US-09-308-935-3
9	68	44.7	19	3	US-09-308-935-15
10	67	44.1	16	3	US-09-308-935-5
11	67	44.1	17	2	US-08-428-131-13
12	67	44.1	17	3	US-09-078-596-13
13	64	42.1	19	3	US-09-308-935-16
14	60	39.5	19	3	US-09-308-935-17
15	56.5	37.2	73	2	US-08-428-131-12
16	56.5	37.2	73	3	US-09-078-596-12
17	56.5	37.2	74	4	US-08-894-139-6
18	55.5	36.5	74	4	US-08-894-139-5
19	55	36.2	14	3	US-09-308-935-11
20	53.5	35.2	74	4	US-08-894-139-7
21	51.5	33.9	75	4	US-08-894-139-9
22	45.5	29.9	69	4	US-08-894-139-8
23	43.5	28.6	100	4	US-08-311-731A-228
24	42	27.6	9	3	US-09-308-935-2
25	42	27.6	61	3	US-08-743-975-9
26	42	27.6	61	4	US-09-263-811-9
27	40	26.3	100	3	US-09-024-024-4

28 39 25.7 75 4 US-09-235-451-16 Sequence 16, Appli  
29 39 25.7 75 4 US-09-235-451-17 Sequence 17, Appli  
30 39 25.7 78 4 US-09-328-352-7148 Sequence 7148, Ap  
31 38 25.0 39 3 US-09-082-279B-1259 Sequence 1259, Ap  
32 38 25.0 39 3 US-09-082-279B-1263 Sequence 1263, Ap  
33 38 25.0 39 4 US-09-315-304B-1259 Sequence 1259, Ap  
34 38 25.0 39 4 US-09-315-304B-1263 Sequence 1263, Ap  
35 38 25.0 39 4 US-09-834-784-1259 Sequence 1259, Ap  
36 38 25.0 39 4 US-09-834-784-1263 Sequence 1263, Ap  
37 37 24.3 87 4 US-09-107-532A-5431 Sequence 5431, Ap  
38 36 23.7 66 4 US-09-205-258-381 Sequence 381, Appl  
39 35.5 23.4 94 1 US-08-241-853-25 Sequence 25, Appl  
40 35.5 23.4 94 2 US-08-850-917-25 Sequence 25, Appl  
41 35 23.0 11 3 US-09-308-935-9 Sequence 9, Appli  
42 35 23.0 26 1 US-07-948-357-3 Sequence 3, Appli  
43 35 23.0 26 2 US-08-450-417-3 Sequence 3, Appli  
44 35 23.0 26 3 US-08-449-741-3 Sequence 3, Appli  
45 35 23.0 26 3 US-08-782-997A-5 Sequence 5, Appli

## ALIGNMENTS

RESULT 1  
US-09-308-935-6  
; Sequence 6, Application US/09308935  
; Patent No. 6268334  
; GENERAL INFORMATION:  
; APPLICANT: La Thangue, Nicholas B  
; APPLICANT: Bandara, Lasantha R  
; TITLE OF INVENTION: Peptide antagonists of DP transcription factors  
; FILE REFERENCE: 620-67  
; CURRENT APPLICATION NUMBER: US/09/308,935  
; CURRENT FILING DATE: 1999-05-27  
; EARLIER APPLICATION NUMBER: PCT/GB97/03506  
; EARLIER FILING DATE: 1997-12-22  
; EARLIER APPLICATION NUMBER: GB 9626589.7  
; EARLIER FILING DATE: 1996-12-20  
; NUMBER OF SEQ ID NOS: 18  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 6  
; LENGTH: 30  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Description of Artificial Sequence: Synthetic  
US-09-308-935-6

Query Match 100.0%; Score 152; DB 3; Length 30;  
Best Local Similarity 100.0%; Pred. NO. 9.4e-17;  
Matches 30; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 YDALNVLMMANNIISKKEIKWIGLPTNSA 30  
Db 1 YDALNVLMMANNIISKKEIKWIGLPTNSA 30

RESULT 2  
US-09-308-935-1  
; Sequence 1, Application US/09308935  
; Patent No. 6268334  
; GENERAL INFORMATION:  
; APPLICANT: La Thangue, Nicholas B  
; APPLICANT: Bandara, Lasantha R  
; TITLE OF INVENTION: Peptide antagonists of DP transcription factors  
; FILE REFERENCE: 620-67  
; CURRENT APPLICATION NUMBER: US/09/308,935  
; CURRENT FILING DATE: 1999-05-27  
; EARLIER APPLICATION NUMBER: PCT/GB97/03506  
; EARLIER FILING DATE: 1997-12-22  
; EARLIER APPLICATION NUMBER: GB 9626589.7  
; EARLIER FILING DATE: 1996-12-20  
; NUMBER OF SEQ ID NOS: 18

```
; SOFTWARE: Patentin Ver. 2.1
; SEQ ID NO 1
; LENGTH: 37
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Synthetic
US-09-308-935-1

Query Match      100.0%; Score 152; DB 3; Length 37;
Best Local Similarity 100.0%; Pred. No. 1.2e-16;
Matches 30; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 YDALNVLMMNNIIISKEKEIKWIGLPTNSA 30
Db 8 YDALNVLMMNNIIISKEKEIKWIGLPTNSA 37

RESULT 3
US-08-428-131-11
; Sequence 11, Application US/08428131
; Patent No. 5863757
; GENERAL INFORMATION:
; APPLICANT: La Thangue, Nicholas Barrie
; TITLE OF INVENTION: Transcription Factor DP-1
; NUMBER OF SEQUENCES: 14
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Nixon & Vanderhye
; STREET: 1100 No. 5863757th Glebe Road, 8th Floor
; CITY: Arlington
; STATE: Virginia
; COUNTRY: U.S.A.
; ZIP: 22201-4714
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25 (EPO)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/428,131
; FILING DATE: 23-JUN-1995
; CLASSIFICATION: 514
; ATTORNEY/AGENT INFORMATION:
; NAME: Arthur R. Crawford
; REGISTRATION NUMBER: 25,327
; REFERENCE/DOCKET NUMBER: 117-181
; TELEPHONE: (703) 816-4000
; TELEFAX: (703) 816-4100
; INFORMATION FOR SEQ ID NO: 11:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 72 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25 (EPO)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/428,131
; FILING DATE: 23-JUN-1995
; CLASSIFICATION: 514
; ATTORNEY/AGENT INFORMATION:
; NAME: Arthur R. Crawford
; REGISTRATION NUMBER: 25,327
; REFERENCE/DOCKET NUMBER: 117-181
; TELEPHONE: (703) 816-4000
; TELEFAX: (703) 816-4100
; INFORMATION FOR SEQ ID NO: 11:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 72 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-428-131-11

Query Match      100.0%; Score 152; DB 2; Length 72;
Best Local Similarity 100.0%; Pred. No. 2.6e-16;
Matches 30; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 YDALNVLMMNNIIISKEKEIKWIGLPTNSA 30
Db 11 YDALNVLMMNNIIISKEKEIKWIGLPTNSA 40

RESULT 4
US-09-078-596-11
; Sequence 11, Application US/09078596
; Patent No. 6150116
; GENERAL INFORMATION:
; APPLICANT: La Thangue, Nicholas Barrie
; TITLE OF INVENTION: Transcription Factor DP-1
; NUMBER OF SEQUENCES: 14
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Nixon & Vanderhye
; STREET: 1100 No. 6150116th Glebe Road, 8th Floor
; CITY: Arlington
; STATE: Virginia
; COUNTRY: U.S.A.
; ZIP: 22201-4714
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/894,139
; FILING DATE: 13-AUG-1997
; CLASSIFICATION: 536
; ATTORNEY/AGENT INFORMATION:
```

```
; NAME: WILSON, MARY J.
; REGISTRATION NUMBER: 32,955
; REFERENCE/DOCKET NUMBER: 620-22
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703) 816-4000
; TELEFAX: (703) 816-4100
; INFORMATION FOR SEQ ID NO: 10:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 74 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; US-08-894-139-10

Query Match      82.9%; Score 126; DB 4; Length 74;
Best Local Similarity 100.0%; Pred. No. 2.7e-12;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 YDALNVLMMNIIISKEKEIKWIGL 25
Db 50 YDALNVLMMNIIISKEKEIKWIGL 74

RESULT 6
US-09-308-935-4
; Sequence 4, Application US/09308935
; Patent No. 6268334
; GENERAL INFORMATION:
; APPLICANT: La Thangue, Nicholas B
; APPLICANT: Bandara, Lasantha R
; TITLE OF INVENTION: Peptide antagonists of DP transcription factors
; FILE REFERENCE: 620-67
; CURRENT APPLICATION NUMBER: US/09/308,935
; CURRENT FILING DATE: 1999-05-27
; EARLIER APPLICATION NUMBER: PCT/GB97/03506
; EARLIER FILING DATE: 1997-12-22
; EARLIER APPLICATION NUMBER: GB 9626589.7
; EARLIER FILING DATE: 1996-12-20
; NUMBER OF SEQ ID NOS: 18
; SOFTWARE: Patentin Ver. 2.1
; SEQ ID NO 4
; LENGTH: 20
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Synthetic
US-09-308-935-4

Query Match      66.4%; Score 101; DB 3; Length 20;
Best Local Similarity 100.0%; Pred. No. 4.3e-09;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 5 NVLMAMNIIISKEKEIKWIG 24
Db 1 NVLMAMNIIISKEKEIKWIG 20

RESULT 7
US-09-308-935-10
; Sequence 10, Application US/09308935
; Patent No. 6268334
; GENERAL INFORMATION:
; APPLICANT: La Thangue, Nicholas B
; APPLICANT: Bandara, Lasantha R
; TITLE OF INVENTION: Peptide antagonists of DP transcription factors
; FILE REFERENCE: 620-67
; CURRENT APPLICATION NUMBER: US/09/308,935
; CURRENT FILING DATE: 1999-05-27
; EARLIER APPLICATION NUMBER: PCT/GB97/03506
; EARLIER FILING DATE: 1997-12-22
; EARLIER APPLICATION NUMBER: GB 9626589.7
; EARLIER FILING DATE: 1996-12-20
```

```
; NUMBER OF SEQ ID NOS: 18
; SOFTWARE: Patentin Ver. 2.1
; SEQ ID NO 10
; LENGTH: 15
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Synthetic
US-09-308-935-10

Query Match      52.6%; Score 80; DB 3; Length 15;
Best Local Similarity 100.0%; Pred. No. 5.3e-06;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 16 EKKEIKWIGLPTNSA 30
Db 1 EKKEIKWIGLPTNSA 15

RESULT 8
US-09-308-935-3
; Sequence 3, Application US/09308935
; Patent No. 6268334
; GENERAL INFORMATION:
; APPLICANT: La Thangue, Nicholas B
; APPLICANT: Bandara, Lasantha R
; TITLE OF INVENTION: Peptide antagonists of DP transcription factors
; FILE REFERENCE: 620-67
; CURRENT APPLICATION NUMBER: US/09/308,935
; CURRENT FILING DATE: 1999-05-27
; EARLIER APPLICATION NUMBER: PCT/GB97/03506
; EARLIER FILING DATE: 1997-12-22
; EARLIER APPLICATION NUMBER: GB 9626589.7
; EARLIER FILING DATE: 1996-12-20
; NUMBER OF SEQ ID NOS: 18
; SOFTWARE: Patentin Ver. 2.1
; SEQ ID NO 3
; LENGTH: 19
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Synthetic
US-09-308-935-3

Query Match      47.4%; Score 72; DB 3; Length 19;
Best Local Similarity 100.0%; Pred. No. 0.00012;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 YDALNVLMMNIIISK 15
Db 5 YDALNVLMMNIIISK 19

RESULT 9
US-09-308-935-15
; Sequence 15, Application US/09308935
; Patent No. 6268334
; GENERAL INFORMATION:
; APPLICANT: La Thangue, Nicholas B
; APPLICANT: Bandara, Lasantha R
; TITLE OF INVENTION: Peptide antagonists of DP transcription factors
; FILE REFERENCE: 620-67
; CURRENT APPLICATION NUMBER: US/09/308,935
; CURRENT FILING DATE: 1999-05-27
; EARLIER APPLICATION NUMBER: PCT/GB97/03506
; EARLIER FILING DATE: 1997-12-22
; EARLIER APPLICATION NUMBER: GB 9626589.7
; EARLIER FILING DATE: 1996-12-20
; NUMBER OF SEQ ID NOS: 18
; SOFTWARE: Patentin Ver. 2.1
; SEQ ID NO 15
; LENGTH: 19
; TYPE: PRT
```

ORGANISM: Artificial Sequence  
FEATURE:  
OTHER INFORMATION: Description of Artificial Sequence: Mutant peptide  
US-09-308-935-15

Query Match 44.7%; Score 68; DB 3; Length 19;  
Best Local Similarity 93.3%; Pred. No. 0.00049; Indels 0; Gaps 0;  
Matches 14; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 YDALNVLMMNIISK 15  
Db 5 YDALNVLMMNIISK 19

RESULT 10  
US-09-308-935-5  
Sequence 5, Application US/09308935  
Patent No. 6268334  
GENERAL INFORMATION:  
APPLICANT: La Thangue, Nicholas B  
APPLICANT: Bandara, Lasantha R  
TITLE OF INVENTION: Peptide antagonists of DP transcription factors  
FILE REFERENCE: 620-67  
CURRENT APPLICATION NUMBER: US/09/308,935  
CURRENT FILING DATE: 1999-05-27  
EARLIER APPLICATION NUMBER: PCT/GB97/03506  
EARLIER FILING DATE: 1997-12-22  
EARLIER APPLICATION NUMBER: GB 9626589.7  
EARLIER FILING DATE: 1996-12-20  
NUMBER OF SEQ ID NOS: 18  
SOFTWARE: PatentIn Ver. 2.1  
SEQ ID NO 5  
LENGTH: 16  
TYPE: PRT  
ORGANISM: Artificial Sequence  
FEATURE:  
OTHER INFORMATION: Description of Artificial Sequence: Synthetic  
US-09-308-935-5

Query Match 44.1%; Score 67; DB 3; Length 16;  
Best Local Similarity 100.0%; Pred. No. 0.00057; Indels 0; Gaps 0;  
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 YDALNVLMMNIIS 14  
Db 3 YDALNVLMMNIIS 16

RESULT 11  
US-08-428-131-13  
Sequence 13, Application US/08428131  
Patent No. 5863757  
GENERAL INFORMATION:  
APPLICANT: La Thangue, Nicholas Barrie  
TITLE OF INVENTION: Transcription Factor DP-1  
NUMBER OF SEQUENCES: 14  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Nixon & Vanderhye  
STREET: 1100 No. 586375th Glebe Road, 8th Floor  
CITY: Arlington  
STATE: Virginia  
COUNTRY: U.S.A.  
ZIP: 22201-4714  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25 (EPO)  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/428,131  
FILING DATE: 23-JUN-1995  
CLASSIFICATION: 514  
ATTORNEY/AGENT INFORMATION:

Query Match 44.1%; Score 67; DB 3; Length 17;  
Best Local Similarity 100.0%; Pred. No. 0.00061; Indels 0; Gaps 0;  
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 YDALNVLMMNIIS 14  
Db 4 YDALNVLMMNIIS 17

NAME: Arthur R. Crawford  
REGISTRATION NUMBER: 25,327  
REFERENCE/DOCKET NUMBER: 117-181  
TELEPHONE: (703) 816-4000  
TELEFAX: (703) 816-4100  
INFORMATION FOR SEQ ID NO: 13:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 17 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-428-131-13

Query Match 44.1%; Score 67; DB 2; Length 17;  
Best Local Similarity 100.0%; Pred. No. 0.00061; Indels 0; Gaps 0;  
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 YDALNVLMMNIIS 14  
Db 4 YDALNVLMMNIIS 17

RESULT 12  
US-09-078-596-13  
Sequence 13, Application US/09078596  
Patent No. 6150116  
GENERAL INFORMATION:  
APPLICANT: La Thangue, Nicholas Barrie  
TITLE OF INVENTION: Transcription Factor DP-1  
NUMBER OF SEQUENCES: 14  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Nixon & Vanderhye  
STREET: 1100 No. 6150116th Glebe Road, 8th Floor  
CITY: Arlington  
STATE: Virginia  
COUNTRY: U.S.A.  
ZIP: 22201-4714  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25 (EPO)  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/078,596  
FILING DATE:  
CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US/08/428,131  
FILING DATE: 23-JUN-1995  
ATTORNEY/AGENT INFORMATION:  
NAME: Arthur R. Crawford  
REGISTRATION NUMBER: 25,327  
REFERENCE/DOCKET NUMBER: 117-181  
TELEPHONE: (703) 816-4000  
TELEFAX: (703) 816-4100  
INFORMATION FOR SEQ ID NO: 13:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 17 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-09-078-596-13

Query Match 44.1%; Score 67; DB 3; Length 17;  
Best Local Similarity 100.0%; Pred. No. 0.00061; Indels 0; Gaps 0;  
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 YDALNVLMMNIIS 14  
Db 4 YDALNVLMMNIIS 17



Db 4 YDALNVLMMNIIIS 17

RESULT 13  
US-09-308-935-16  
; Sequence 16, Application US/09308935  
; Patent No. 6268334  
; GENERAL INFORMATION:  
; APPLICANT: La Thangue, Nicholas B  
; APPLICANT: Bandara, Lasantha R  
; TITLE OF INVENTION: Peptide antagonists of DP transcription factors  
; FILE REFERENCE: 620-67  
; CURRENT APPLICATION NUMBER: US/09/308,935  
; CURRENT FILING DATE: 1999-05-27  
; EARLIER APPLICATION NUMBER: PCT/GB97/03506  
; EARLIER FILING DATE: 1997-12-22  
; EARLIER APPLICATION NUMBER: GB 9626589.7  
; EARLIER FILING DATE: 1996-12-20  
; NUMBER OF SEQ ID NOS: 18  
; SOFTWARE: Patentin Ver. 2.1  
; SEQ ID NO 16  
; LENGTH: 19  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Description of Artificial Sequence: Mutant peptide  
US-09-308-935-16

Query Match 42.18; Score 64; DB 3; Length 19;  
Best Local Similarity 93.38; Pred. No. 0.002; Mismatches 0; Indels 1; Gaps 0;  
Matches 14; Conservative 0; Mismatches 1; Indels 1; Gaps 0;

Qy 1 YDALNVLMMNIIISK 15  
| | | | | | | | | | | | | | | | | | | | | |  
Db 5 YDALNVLMMNIIISK 19

RESULT 14  
US-09-308-935-17  
; Sequence 17, Application US/09308935  
; Patent No. 6268334  
; GENERAL INFORMATION:  
; APPLICANT: La Thangue, Nicholas B  
; APPLICANT: Bandara, Lasantha R  
; TITLE OF INVENTION: Peptide antagonists of DP transcription factors  
; FILE REFERENCE: 620-67  
; CURRENT APPLICATION NUMBER: US/09/308,935  
; CURRENT FILING DATE: 1999-05-27  
; EARLIER APPLICATION NUMBER: PCT/GB97/03506  
; EARLIER FILING DATE: 1997-12-22  
; EARLIER APPLICATION NUMBER: GB 9626589.7  
; EARLIER FILING DATE: 1996-12-20  
; NUMBER OF SEQ ID NOS: 18  
; SOFTWARE: Patentin Ver. 2.1  
; SEQ ID NO 17  
; LENGTH: 19  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Description of Artificial Sequence: Mutant peptide  
US-09-308-935-17

Query Match 39.58; Score 60; DB 3; Length 19;  
Best Local Similarity 86.78; Pred. No. 0.0083; Mismatches 0; Indels 2; Gaps 0;  
Matches 13; Conservative 0; Mismatches 2; Indels 2; Gaps 0;

Qy 1 YDALNVLMMNIIISK 15  
| | | | | | | | | | | | | | | | | | | | | |  
Db 5 YDALNVLMMNIIISK 19

RESULT 15  
US-08-428-131-12

Sequence 12, Application US/08428131  
; Patent No. 5863757  
; GENERAL INFORMATION:  
; APPLICANT: La Thangue, Nicholas Barrie  
; TITLE OF INVENTION: Transcription Factor DP-1  
; NUMBER OF SEQUENCES: 14  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Nixon & Vanderhye  
; STREET: 1100 No. 5863757th Glebe Road, 8th Floor  
; CITY: Arlington  
; STATE: Virginia  
; COUNTRY: U.S.A.  
; ZIP: 22201-4714  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patentin Release #1.0, Version #1.25 (EPO)  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/428,131  
; FILING DATE: 23-JUN-1995  
; CLASSIFICATION: 514  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Arthur R. Crawford  
; REGISTRATION NUMBER: 25,327  
; REFERENCE/DOCKET NUMBER: 117-181  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (703) 816-4000  
; TELEFAX: (703) 816-4100  
; INFORMATION FOR SEQ ID NO: 12:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 73 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
US-08-428-131-12

Query Match 37.28; Score 56.5; DB 2; Length 73;  
Best Local Similarity 42.98; Pred. No. 0.14; Mismatches 12; Conservative 6; Mismatches 9; Indels 1; Gaps 1;  
Matches 12; Conservative 6; Mismatches 9; Indels 1; Gaps 1;

Qy 1 YDALNVLMMNIIISKE-KKEIKWIGLPT 27  
| | | | | | | | | | | | | | | | | | | | | |  
Db 11 YDITNVLEGIQIAKSKNHIQWLGSH 38

RESULT 16  
US-09-078-596-12  
; Sequence 12, Application US/09078596  
; Patent No. 6150116  
; GENERAL INFORMATION:  
; APPLICANT: La Thangue, Nicholas Barrie  
; TITLE OF INVENTION: Transcription Factor DP-1  
; NUMBER OF SEQUENCES: 14  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Nixon & Vanderhye  
; STREET: 1100 No. 6150116th Glebe Road, 8th Floor  
; CITY: Arlington  
; STATE: Virginia  
; COUNTRY: U.S.A.  
; ZIP: 22201-4714  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patentin Release #1.0, Version #1.25 (EPO)  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/078,596  
; FILING DATE:  
; CLASSIFICATION:  
; PRIOR APPLICATION NUMBER: US/08/428,131

us-09-900-147-6.ra1

Tue Feb 17 11:55:54 2004

```
; FILING DATE: 23-JUN-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Arthur R. Crawford
; REGISTRATION NUMBER: 25,327
; REFERENCE/DOCKET NUMBER: 117-181
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703) 816-4000
; TELEFAX: (703) 816-4100
; INFORMATION FOR SEQ ID NO: 12:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 73 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-09-078-596-12

Query Match 37.2%; Score 56.5; DB 3; Length 73;
Best Local Similarity 42.9%; Pred. No. 0.14;
Matches 12; Conservative 6; Mismatches 9; Indels 1; Gaps 1;

Qy 1 YDALNVLMMNIIISKE-KKEIKWIGLPT 27
Db 11 YDITNVLEGIQLIRKKKKNHIQWLG 38

RESULT 17
US-08-894-139-6
; Sequence 6, Application US/08894139
; Patent No. 6448376
; GENERAL INFORMATION:
; APPLICANT: LA THANGUE, NICHOLAS B.
; APPLICANT: BERNARDS, RENE
; APPLICANT: HIJMAN, ELEANORE M.
; TITLE OF INVENTION: TRANSCRIPTION FACTOR E2F-5
; NUMBER OF SEQUENCES: 25
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: NIXON & VANDERHUYE P.C.
; STREET: 1100 NORTH GLEBE ROAD
; CITY: ARLINGTON
; STATE: VIRGINIA
; COUNTRY: U.S.A.
; ZIP: 22201-4714
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/894,139
; FILING DATE: 13-AUG-1997
; CLASSIFICATION: 536
; ATTORNEY/AGENT INFORMATION:
; NAME: WILSON, MARY J.
; REGISTRATION NUMBER: 32,955
; REFERENCE/DOCKET NUMBER: 620-22
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703) 816-4000
; TELEFAX: (703) 816-4100
; INFORMATION FOR SEQ ID NO: 5:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 74 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; US-08-894-139-5

Query Match 36.5%; Score 55.5; DB 4; Length 74;
Best Local Similarity 44.0%; Pred. No. 0.2;
Matches 11; Conservative 6; Mismatches 7; Indels 1; Gaps 1;

Qy 1 YDALNVLMMNIIISKE-KKEIKWIG 24
Db 49 YDITNVLEGIQLIRKKKKNHIQWLG 73

RESULT 19
US-09-308-935-11
; Sequence 11, Application US/09308935
; Patent No. 6268334
; GENERAL INFORMATION:
; APPLICANT: La Thangue, Nicholas B
; APPLICANT: Bandara, Lisantha R
; TITLE OF INVENTION: Peptide antagonists of DP transcription factors
; FILE REFERENCE: 620-67
; CURRENT APPLICATION NUMBER: US/09/308,935
; CURRENT FILING DATE: 1999-05-27
; EARLIER APPLICATION NUMBER: PCT/GB97/03506
; EARLIER FILING DATE: 1997-12-22
; EARLIER APPLICATION NUMBER: GB 9626589.7
; EARLIER FILING DATE: 1996-12-20
; NUMBER OF SEQ ID NOS: 18
; SOFTWARE: Patent In Ver. 2.1
; SEQ ID NO 11
```

```
; LENGTH: 14
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Synthetic
US-09-308-935-11
Query Match 36.2%; Score 55; DB 3; Length 14;
Best Local Similarity 100.0%; Pred. No. 0.035;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 YDALNVLMMANN 11
Db 4 YDALNVLMMANN 14

RESULT 20
US-08-894-139-7
; Sequence 7, Application US/08894139
; Patent No. 6448376
; GENERAL INFORMATION:
; APPLICANT: LA THANGUE, NICHOLAS B.
; APPLICANT: BERNARDS, RENE
; APPLICANT: HIJMAN, ELEANORE M.
; TITLE OF INVENTION: TRANSCRIPTION FACTOR E2F-5
; NUMBER OF SEQUENCES: 25
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: NIXON & VANDERHYE P.C.
; STREET: 1100 NORTH GLEBE ROAD
; CITY: ARLINGTON
; STATE: VIRGINIA
; COUNTRY: U.S.A.
; ZIP: 22201-4714
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/894,139
; FILING DATE: 13-AUG-1997
; CLASSIFICATION: 536
; ATTORNEY/AGENT INFORMATION:
; NAME: WILSON, MARY J.
; REGISTRATION NUMBER: 32,955
; REFERENCE/DOCKET NUMBER: 620-22
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703) 816-4000
; TELEFAX: (703) 816-4100
; INFORMATION FOR SEQ ID NO: 9:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 75 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-08-894-139-9
Query Match 33.9%; Score 51.5; DB 4; Length 75;
Best Local Similarity 42.3%; Pred. No. 0.83;
Matches 11; Conservative 6; Mismatches 8; Indels 1; Gaps 1;

QY 1 YDALNVLMMANNIISKE-KKEIKWIGL 25
Db 50 YDITNVLEGIDLIEKSKNSIQWKG 75

RESULT 22
US-08-894-139-8
; Sequence 8, Application US/08894139
; Patent No. 6448376
; GENERAL INFORMATION:
; APPLICANT: LA THANGUE, NICHOLAS B.
; APPLICANT: BERNARDS, RENE
; APPLICANT: HIJMAN, ELEANORE M.
; TITLE OF INVENTION: TRANSCRIPTION FACTOR E2F-5
; NUMBER OF SEQUENCES: 25
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: NIXON & VANDERHYE P.C.
; STREET: 1100 NORTH GLEBE ROAD
; CITY: ARLINGTON
; STATE: VIRGINIA
; COUNTRY: U.S.A.
; ZIP: 22201-4714
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/894,139
; FILING DATE: 13-AUG-1997
; CLASSIFICATION: 536
; ATTORNEY/AGENT INFORMATION:
; NAME: WILSON, MARY J.
; REGISTRATION NUMBER: 32,955
; REFERENCE/DOCKET NUMBER: 620-22
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703) 816-4000
; TELEFAX: (703) 816-4100
; INFORMATION FOR SEQ ID NO: 7:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 74 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-08-894-139-7
Query Match 35.2%; Score 53.5; DB 4; Length 74;
Best Local Similarity 40.0%; Pred. No. 0.4;
Matches 10; Conservative 7; Mismatches 7; Indels 1; Gaps 1;

QY 1 YDALNVLMMANNIISKE-KKEIKWIG 24
Db 49 YDITNVLEGIDLIEKSKNSIQWKG 73

RESULT 21
US-08-894-139-9
; Sequence 9, Application US/08894139
; Patent No. 6448376
; GENERAL INFORMATION:
; APPLICANT: LA THANGUE, NICHOLAS B.
```

REGISTRATION NUMBER: 32,955  
REFERENCE/DOCKET NUMBER: 620-22  
TELEPHONE: (703) 816-4000  
TELEFAX: (703) 816-4100  
INFORMATION FOR SEQ ID NO: 8:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 69 amino acids  
TYPE: amino acid  
STRANDEDNESS:  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
US-08-894-139-8

Query Match 29.9%; Score 45.5; DB 4; Length 69;  
Best Local Similarity 38.5%; Pred. No. 6.3;  
Matches 10; Conservative 5; Mismatches 10; Indels 1; Gaps 1;

QY 1 YDALNVLMMNNIIISKE-KKEIKWGL 25  
DB 44 YDITNVLEGIGLIEKSKNSTQWRGV 69

RESULT 23  
US-08-311-731A-228  
; Sequence 228, Application US/08311731A  
; Patent No. 6583266  
; GENERAL INFORMATION:  
; APPLICANT: SMITH, DOUGLAS  
; APPLICANT: MAO, JEN-I  
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES  
; TITLE OF INVENTION: RELATING TO MYCOBACTERIUM TUBERCULOSIS AND LAPRAE FOR  
; TITLE OF INVENTION: DIAGNOSTICS AND THERAPEUTICS  
; NUMBER OF SEQUENCES: 411  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: WOLF, GREENFIELD & SACKS, P.C.  
; STREET: 600 ATLANTIC AVENUE  
; CITY: BOSTON  
; STATE: MASSACHUSETTS  
; COUNTRY: USA  
; ZIP: 02210  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/311,731A  
; FILING DATE:  
; CLASSIFICATION: 530  
; ATTORNEY/AGENT INFORMATION:  
; NAME: GATES, EDWARD R.  
; REGISTRATION NUMBER: 31,616  
; REFERENCE/DOCKET NUMBER: C0044/7125  
; TELEPHONE: 617/720-3500  
; TELEFAX: 617/720-2441  
; INFORMATION FOR SEQ ID NO: 228:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 100 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
; HYPOTHETICAL: YES  
; ORIGINAL SOURCE:  
; ORGANISM: Mycobacterium leprae  
US-08-311-731A-228

Query Match 28.6%; Score 43.5; DB 4; Length 100;  
Best Local Similarity 37.0%; Pred. No. 20;  
Matches 10; Conservative 6; Mismatches 8; Indels 3; Gaps 1;

QY 3 ALNVLMMNNIIISKEKKEIKWGLPTNS 29

Db 14 ACADVAMLM---SRIESEASWMLPVD 37

RESULT 24  
US-09-308-935-2  
; Sequence 2, Application US/09308935  
; Patent No. 6268334  
; GENERAL INFORMATION:  
; APPLICANT: La Thangue, Nicholas B  
; APPLICANT: Bandara, Laseantha R  
; TITLE OF INVENTION: Peptide antagonists of DP transcription factors  
; FILE REFERENCE: 620-67  
; CURRENT APPLICATION NUMBER: US/09/308,935  
; CURRENT FILING DATE: 1999-05-27  
; EARLIER APPLICATION NUMBER: PCT/GB97/03506  
; EARLIER FILING DATE: 1997-12-22  
; EARLIER APPLICATION NUMBER: GB 9626589.7  
; EARLIER FILING DATE: 1996-12-20  
; NUMBER OF SEQ ID NOS: 18  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 2  
; LENGTH: 9  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Description of Artificial Sequence: Synthetic  
US-09-308-935-2

Query Match 27.6%; Score 42; DB 3; Length 9;  
Best Local Similarity 100.0%; Pred. No. 2.5e+05;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 5 NVLMAMNNII 13  
DB 1 NVLMAMNNII 9

RESULT 25  
US-08-743-975-9  
; Sequence 9, Application US/08743975  
; Patent No. 6057434  
; GENERAL INFORMATION:  
; APPLICANT: Jian Ni, Guo-Liang Yu and Reiner Gentz  
; TITLE OF INVENTION: Mammary Transforming Protein  
; NUMBER OF SEQUENCES: 9  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: CARELLA, BYRNE, BAIN, GILFILLAN,  
; ADDRESSEE: CECCHI, STEWART & OLSTEIN  
; STREET: 6 BECKER FARM ROAD  
; CITY: ROSELAND  
; STATE: NEW JERSEY  
; COUNTRY: USA  
; ZIP: 07068  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: 3.5 INCH DISKETTE  
; COMPUTER: IBM PS/2  
; OPERATING SYSTEM: MS-DOS  
; SOFTWARE: WORD PERFECT 5.1  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/743,975  
; FILING DATE: 01 NOVEMBER 1996  
; CLASSIFICATION: 514  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 60/006,187  
; FILING DATE: 02 NOVEMBER 1995  
; ATTORNEY/AGENT INFORMATION:  
; NAME: MULLINS, J.G.  
; REGISTRATION NUMBER: 33,073  
; REFERENCE/DOCKET NUMBER: 325800-507 (PF212)  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 201-994-1700  
; TELEFAX: 201-994-1744



```

; TYPE: PRT
; ORGANISM: Caliphora
US-09-235-451-16

Query Match      25.7%; Score 39; DB 4; Length 75;
Best Local Similarity 18.8%; Pred. No. 70;
Matches 6; Conservative 10; Mismatches 6; Indels 10; Gaps 1;

QY 1 YDALNVLMMNMI-----ISKEKKEIKW 22
Db 35 YSVINIIVLLNMLIAMNSYQIISERADVEW 66

RESULT 29
US-09-235-451-17
; Sequence 17, Application US/09235451
; GENERAL INFORMATION:
; APPLICANT: Julius, David J.
; APPLICANT: Caterina, Michael J.
; APPLICANT: Brake, Anthony J.
; TITLE OF INVENTION: NUCLEIC ACID SEQUENCES ENCODING
; TITLE OF INVENTION: CAPSAICIN RECEPTOR AND CAPSAICIN RECEPTOR-RELATED
; TITLE OF INVENTION: POLYPEPTIDES AND USES THEREOF
; FILE REFERENCE: 9076/084CIP
; CURRENT APPLICATION NUMBER: US/09/235,451
; PRIOR FILING DATE: 1999-01-22
; PRIOR FILING DATE: 1998-01-22
; PRIOR FILING DATE: 1997-08-20
; NUMBER OF SEQ ID NOS: 48
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 17
; LENGTH: 75
; TYPE: PRT
; ORGANISM: D. melanogaster
US-09-235-451-17

Query Match      25.7%; Score 39; DB 4; Length 75;
Best Local Similarity 18.8%; Pred. No. 70;
Matches 6; Conservative 10; Mismatches 6; Indels 10; Gaps 1;

QY 1 YDALNVLMMNMI-----ISKEKKEIKW 22
Db 35 YSVINIIVLLNMLIAMNSYQIISERADVEW 66

RESULT 30
US-09-328-352-7148
; Sequence 7148, Application US/09328352
; Patent No. 6562958
; GENERAL INFORMATION:
; APPLICANT: Gary L. Breton et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO ACINETOBACTER
; TITLE OF INVENTION: BAUMANNII FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: GTC99-03PA
; CURRENT APPLICATION NUMBER: US/09/328,352
; CURRENT FILING DATE: 1999-06-04
; NUMBER OF SEQ ID NOS: 8252
; SEQ ID NO 7148
; LENGTH: 78
; TYPE: PRT
; ORGANISM: Acinetobacter baumannii
US-09-328-352-7148

Query Match      25.7%; Score 39; DB 4; Length 78;
Best Local Similarity 31.2%; Pred. No. 73;
Matches 5; Conservative 8; Mismatches 3; Indels 0; Gaps 0;

QY 8 MAMNIIISKEKKEIKWI 23
Db 40 VAMSLTRPRKALDWL 55

; TYPE: PRT
; ORGANISM: Caliphora
US-09-235-451-16

Query Match      25.7%; Score 39; DB 4; Length 75;
Best Local Similarity 18.8%; Pred. No. 70;
Matches 6; Conservative 10; Mismatches 6; Indels 10; Gaps 1;

QY 1 YDALNVLMMNMI-----ISKEKKEIKW 22
Db 35 YSVINIIVLLNMLIAMNSYQIISERADVEW 66

RESULT 29
US-09-235-451-17
; Sequence 17, Application US/09235451
; GENERAL INFORMATION:
; APPLICANT: Julius, David J.
; APPLICANT: Caterina, Michael J.
; APPLICANT: Brake, Anthony J.
; TITLE OF INVENTION: NUCLEIC ACID SEQUENCES ENCODING
; TITLE OF INVENTION: CAPSAICIN RECEPTOR AND CAPSAICIN RECEPTOR-RELATED
; TITLE OF INVENTION: POLYPEPTIDES AND USES THEREOF
; FILE REFERENCE: 9076/084CIP
; CURRENT APPLICATION NUMBER: US/09/235,451
; PRIOR FILING DATE: 1999-01-22
; PRIOR FILING DATE: 1998-01-22
; PRIOR FILING DATE: 1997-08-20
; NUMBER OF SEQ ID NOS: 48
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 17
; LENGTH: 75
; TYPE: PRT
; ORGANISM: D. melanogaster
US-09-235-451-17

Query Match      25.7%; Score 39; DB 4; Length 75;
Best Local Similarity 18.8%; Pred. No. 70;
Matches 6; Conservative 10; Mismatches 6; Indels 10; Gaps 1;

QY 1 YDALNVLMMNMI-----ISKEKKEIKW 22
Db 35 YSVINIIVLLNMLIAMNSYQIISERADVEW 66

RESULT 30
US-09-328-352-7148
; Sequence 7148, Application US/09328352
; Patent No. 6562958
; GENERAL INFORMATION:
; APPLICANT: Gary L. Breton et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO ACINETOBACTER
; TITLE OF INVENTION: BAUMANNII FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: GTC99-03PA
; CURRENT APPLICATION NUMBER: US/09/328,352
; CURRENT FILING DATE: 1999-06-04
; NUMBER OF SEQ ID NOS: 8252
; SEQ ID NO 7148
; LENGTH: 78
; TYPE: PRT
; ORGANISM: Acinetobacter baumannii
US-09-328-352-7148

Query Match      25.7%; Score 39; DB 4; Length 78;
Best Local Similarity 31.2%; Pred. No. 73;
Matches 5; Conservative 8; Mismatches 3; Indels 0; Gaps 0;

QY 8 MAMNIIISKEKKEIKWI 23
Db 40 VAMSLTRPRKALDWL 55
```

```

; TYPE: PRT
; ORGANISM: Caliphora
US-09-082-279B-1259

Query Match      25.0%; Score 38; DB 3; Length 39;
Best Local Similarity 35.7%; Pred. No. 47;
Matches 10; Conservative 3; Mismatches 9; Indels 6; Gaps 1;

QY 4 LNVLMAMNIIISKEKKE-----IKWIGL 25
Db 8 ITALLEQAQIQQEKNEYELQKLIKWAGL 35

RESULT 32
US-09-082-279B-1263
; Sequence 1263, Application US/09082279B
; Patent No. 6258782
; GENERAL INFORMATION:
; APPLICANT: Barney, Shawn
; APPLICANT: Guthrie, Kelly
; APPLICANT: Merutka, Gene
; APPLICANT: Anwer, Mohamed
; APPLICANT: Lambert, Dennis
; TITLE OF INVENTION: HYBRID POLYPEPTIDES WITH ENHANCED
; TITLE OF INVENTION: PHARMACOKINETIC PROPERTIES
; FILE REFERENCE: 7872-043
; CURRENT APPLICATION NUMBER: US/09/082,279B
; CURRENT FILING DATE: 1998-05-20
; NUMBER OF SEQ ID NOS: 1515
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 1263
; LENGTH: 39
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Core polypeptide
US-09-082-279B-1263

Query Match      25.0%; Score 38; DB 3; Length 39;
Best Local Similarity 35.7%; Pred. No. 47;
Matches 10; Conservative 3; Mismatches 9; Indels 6; Gaps 1;

QY 4 LNVLMAMNIIISKEKKE-----IKWIGL 25
Db 8 ITALLEQAQIQQEKNEYELQKLIKWAGL 35

RESULT 33
US-09-315-304B-1259
; Sequence 1259, Application US/09315304B
; Patent No. 6348568
```

```
; GENERAL INFORMATION:
; APPLICANT: Barney, S.
; APPLICANT: Guthrie, K.
; APPLICANT: Merutka, G.
; APPLICANT: Anwer, M.
; APPLICANT: Lambert, D.
; TITLE OF INVENTION: HYBRID POLYPEPTIDES WITH ENHANCED PHARMACOKINETIC
; FILE REFERENCE: 7872-052
; CURRENT APPLICATION NUMBER: US/09/315,304B
; CURRENT FILING DATE: 1999-05-20
; PRIOR APPLICATION NUMBER: 09/082,279
; PRIOR FILING DATE: 1998-05-20
; NUMBER OF SEQ ID NOS: 1667
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 1259
; LENGTH: 39
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Core polypeptide
US-09-315-304B-1259

Query Match      25.0%; Score 38; DB 4; Length 39;
Best Local Similarity 35.7%; Pred. No. 47;
Matches 10; Conservative 3; Mismatches 9; Indels 9; Gaps 1;

QY      4 LNVLMAMNIISKKE-----IKWIGL 25
      : : : : : : : : : : : :
Db      8 ITALLEQAQIQQEKNEYELQKLIKWAGL 35

RESULT 34
US-09-315-304B-1263
; Sequence 1263, Application US/09315304B
; Patent No. 6348568
; GENERAL INFORMATION:
; APPLICANT: Barney, S.
; APPLICANT: Guthrie, K.
; APPLICANT: Merutka, G.
; APPLICANT: Anwer, M.
; APPLICANT: Lambert, D.
; TITLE OF INVENTION: HYBRID POLYPEPTIDES WITH ENHANCED PHARMACOKINETIC
; FILE REFERENCE: 7872-052
; CURRENT APPLICATION NUMBER: US/09/315,304B
; CURRENT FILING DATE: 1999-05-20
; PRIOR APPLICATION NUMBER: 09/082,279
; PRIOR FILING DATE: 1998-05-20
; NUMBER OF SEQ ID NOS: 1667
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 1263
; LENGTH: 39
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Core polypeptide
US-09-315-304B-1263

Query Match      25.0%; Score 38; DB 4; Length 39;
Best Local Similarity 35.7%; Pred. No. 47;
Matches 10; Conservative 3; Mismatches 9; Indels 9; Gaps 1;

QY      4 LNVLMAMNIISKKE-----IKWIGL 25
      : : : : : : : : : : : :
Db      8 ITALLEQAQIQQEKNEYELQKLIKWAGL 35

RESULT 35
US-09-834-784-1259
; Sequence 1259, Application US/09834784
; Patent No. 6562787
; GENERAL INFORMATION:
; APPLICANT: Barney, Shawn
; APPLICANT: Guthrie, Kelly
; APPLICANT: Merutka, Gene
; APPLICANT: Anwer, Mohamed
; APPLICANT: Lambert, Dennis
; TITLE OF INVENTION: HYBRID POLYPEPTIDES WITH ENHANCED
; FILE REFERENCE: 7872-043
; CURRENT APPLICATION NUMBER: US/09/834,784
; CURRENT FILING DATE: 2001-04-13
; PRIOR APPLICATION NUMBER: 09/082,279
; PRIOR FILING DATE: 1998-05-20
; NUMBER OF SEQ ID NOS: 1515
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 1259
; LENGTH: 39
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Core polypeptide
US-09-834-784-1259

Query Match      25.0%; Score 38; DB 4; Length 39;
Best Local Similarity 35.7%; Pred. No. 47;
Matches 10; Conservative 3; Mismatches 9; Indels 9; Gaps 1;

QY      4 LNVLMAMNIISKKE-----IKWIGL 25
      : : : : : : : : : : : :
Db      8 ITALLEQAQIQQEKNEYELQKLIKWAGL 35

RESULT 36
US-09-834-784-1263
; Sequence 1263, Application US/09834784
; Patent No. 6562787
; GENERAL INFORMATION:
; APPLICANT: Barney, Shawn
; APPLICANT: Guthrie, Kelly
; APPLICANT: Merutka, Gene
; APPLICANT: Anwer, Mohamed
; APPLICANT: Lambert, Dennis
; TITLE OF INVENTION: HYBRID POLYPEPTIDES WITH ENHANCED
; FILE REFERENCE: 7872-043
; CURRENT APPLICATION NUMBER: US/09/834,784
; CURRENT FILING DATE: 2001-04-13
; PRIOR APPLICATION NUMBER: 09/082,279
; PRIOR FILING DATE: 1998-05-20
; NUMBER OF SEQ ID NOS: 1515
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 1263
; LENGTH: 39
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Core polypeptide
US-09-834-784-1263

Query Match      25.0%; Score 38; DB 4; Length 39;
Best Local Similarity 35.7%; Pred. No. 47;
Matches 10; Conservative 3; Mismatches 9; Indels 9; Gaps 1;

QY      4 LNVLMAMNIISKKE-----IKWIGL 25
      : : : : : : : : : : : :
Db      8 ITALLEQAQIQQEKNEYELQKLIKWAGL 35

RESULT 37
US-09-107-532A-5431
; Sequence 5431, Application US/09107532A
; Patent No. 6583275
; GENERAL INFORMATION:
; APPLICANT: Lynn A Doucette-Stamm and David Bush
```

TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO  
ENTEROCOCCUS FAECIUM FOR DIAGNOSTICS AND THERAPEUTICS

NUMBER OF SEQUENCES: 7310  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: GENOME THERAPEUTICS CORPORATION  
STREET: 100 Beaver Street  
CITY: Waltham  
STATE: Massachusetts  
COUNTRY: USA  
ZIP: 02354

COMPUTER READABLE FORM:  
MEDIUM TYPE: CD/ROM ISO9660  
COMPUTER: PC  
OPERATING SYSTEM: <Unknown>  
SOFTWARE: ASCII

CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/107,532A  
FILING DATE: 30-Jun-1998

PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 60/085,598  
FILING DATE: 14 May 1998  
APPLICATION NUMBER: 60/051571  
FILING DATE: July 2, 1997

ATTORNEY/AGENT INFORMATION:  
NAME: Ariniello, Pamela Deneke  
REGISTRATION NUMBER: 40,489  
REFERENCE/DOCKET NUMBER: GTC-012  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (781)893-5007  
TELEFAX: (781)893-8277

INFORMATION FOR SEQ ID NO: 5431:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 87 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
HYPOTHETICAL: YES  
ORIGINAL SOURCE:  
ORGANISM: Enterococcus faecium  
FEATURE:  
NAME/KEY: misc feature  
LOCATION: (B) LOCATION 1...87  
SEQUENCE DESCRIPTION: SEQ ID NO: 5431:

US-09-107-532A-5431

Query Match 24.3%; Score 37; DB 4; Length 87;  
Best Local Similarity 46.7%; Pred. No. 1.7e+02;  
Matches 7; Conservative 5; Mismatches 3; Indels 0; Gaps 0;

Qy 7 LMANNIISKKEIK 21  
|:::|:::|:::|  
Db 23 LLAIQLSKKKKK 37

RESULT 38  
US-09-205-258-381  
Sequence 381, Application US/09205258  
Patent No. 6525174  
GENERAL INFORMATION:  
APPLICANT: Young et al.  
TITLE OF INVENTION: 207 Human Secreted Proteins  
FILE REFERENCE: P2007P1  
CURRENT APPLICATION NUMBER: US/09/205,258  
CURRENT FILING DATE: 1998-12-04  
EARLIER APPLICATION NUMBER: PCT/US98/11422  
EARLIER FILING DATE: 1998-06-04  
EARLIER APPLICATION NUMBER: 60/048,885  
EARLIER FILING DATE: 1997-06-06  
EARLIER APPLICATION NUMBER: 60/049,375  
EARLIER FILING DATE: 1997-06-06  
EARLIER APPLICATION NUMBER: 60/048,881  
EARLIER FILING DATE: 1997-06-06  
EARLIER APPLICATION NUMBER: 60/048,880

EARLIER FILING DATE: 1997-06-06  
EARLIER APPLICATION NUMBER: 60/048,896  
EARLIER FILING DATE: 1997-06-06  
EARLIER APPLICATION NUMBER: 60/049,020  
EARLIER FILING DATE: 1997-06-06  
EARLIER APPLICATION NUMBER: 60/048,876  
EARLIER FILING DATE: 1997-06-06  
EARLIER APPLICATION NUMBER: 60/048,895  
EARLIER FILING DATE: 1997-06-06  
EARLIER APPLICATION NUMBER: 60/048,884  
EARLIER FILING DATE: 1997-06-06  
EARLIER APPLICATION NUMBER: 60/048,894  
EARLIER FILING DATE: 1997-06-06  
EARLIER APPLICATION NUMBER: 60/048,971  
EARLIER FILING DATE: 1997-06-06  
EARLIER APPLICATION NUMBER: 60/048,964  
EARLIER FILING DATE: 1997-06-06  
EARLIER APPLICATION NUMBER: 60/048,882  
EARLIER FILING DATE: 1997-06-06  
EARLIER APPLICATION NUMBER: 60/048,899  
EARLIER FILING DATE: 1997-06-06  
EARLIER APPLICATION NUMBER: 60/048,893  
EARLIER FILING DATE: 1997-06-06  
EARLIER APPLICATION NUMBER: 60/048,900  
EARLIER FILING DATE: 1997-06-06  
EARLIER APPLICATION NUMBER: 60/048,901  
EARLIER FILING DATE: 1997-06-06  
EARLIER APPLICATION NUMBER: 60/048,892  
EARLIER FILING DATE: 1997-06-06  
EARLIER APPLICATION NUMBER: 60/048,915  
EARLIER FILING DATE: 1997-06-06  
EARLIER APPLICATION NUMBER: 60/049,019  
EARLIER FILING DATE: 1997-06-06  
EARLIER APPLICATION NUMBER: 60/048,970  
EARLIER FILING DATE: 1997-06-06  
EARLIER APPLICATION NUMBER: 60/048,972  
EARLIER FILING DATE: 1997-06-06  
EARLIER APPLICATION NUMBER: 60/048,916  
EARLIER FILING DATE: 1997-06-06  
EARLIER APPLICATION NUMBER: 60/049,373  
EARLIER FILING DATE: 1997-06-06  
EARLIER APPLICATION NUMBER: 60/048,875  
EARLIER FILING DATE: 1997-06-06  
EARLIER APPLICATION NUMBER: 60/049,374  
EARLIER FILING DATE: 1997-06-06  
EARLIER APPLICATION NUMBER: 60/048,917  
EARLIER FILING DATE: 1997-06-06  
EARLIER APPLICATION NUMBER: 60/048,949  
EARLIER FILING DATE: 1997-06-06  
EARLIER APPLICATION NUMBER: 60/048,974  
EARLIER FILING DATE: 1997-06-06  
EARLIER APPLICATION NUMBER: 60/048,883  
EARLIER FILING DATE: 1997-06-06  
EARLIER APPLICATION NUMBER: 60/048,897  
EARLIER FILING DATE: 1997-06-06  
EARLIER APPLICATION NUMBER: 60/048,898  
EARLIER FILING DATE: 1997-06-06  
EARLIER APPLICATION NUMBER: 60/048,962  
EARLIER FILING DATE: 1997-06-06  
EARLIER APPLICATION NUMBER: 60/048,963  
EARLIER FILING DATE: 1997-06-06  
EARLIER APPLICATION NUMBER: 60/048,877  
EARLIER FILING DATE: 1997-06-06  
EARLIER APPLICATION NUMBER: 60/048,878  
EARLIER FILING DATE: 1997-06-06  
EARLIER APPLICATION NUMBER: 60/070,923  
EARLIER FILING DATE: 1997-12-18  
EARLIER APPLICATION NUMBER: 60/092,921  
EARLIER FILING DATE: 1998-07-15  
EARLIER APPLICATION NUMBER: 60/094,657  
EARLIER FILING DATE: 1998-07-30  
NUMBER OF SEQ ID NOS: 1227  
SOFTWARE: PatentIn Ver. 2.0



SEQ ID NO 381  
LENGTH: 66  
TYPE: PRT  
ORGANISM: Homo sapiens  
FEATURE:  
NAME/KEY: SITE  
LOCATION: (14)  
OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids  
FEATURE:  
NAME/KEY: SITE  
LOCATION: (62)  
OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids  
US-09-205-258-381

Query Match 23.4%; Score 36; DB 4; Length 66;  
Best Local Similarity 31.6%; Pred. No. 1.8e+02;  
Matches 6; Conservative 8; Mismatches 5; Indels 0; Gaps 0;

QY 3 ALNVLAMNIIISKEKKEIK 21  
DB 36 SINLSLLTLIRKKKKKK 54

RESULT 39  
US-08-241-853-25  
Sequence 25, Application US/08241853  
Patent No. 5693488  
GENERAL INFORMATION:  
APPLICANT: Fang, Kathy S.  
APPLICANT: Hanafusa, Hidesaburo  
TITLE OF INVENTION: A TRANSMEMBRANE TYROSINE PHOSPHATASE  
NUMBER OF SEQUENCES: 37  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Klauber & Jackson  
STREET: 411 Hackensack Avenue  
CITY: Hackensack  
STATE: New Jersey  
COUNTRY: USA  
ZIP: 07601  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent In Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/241,853  
FILING DATE: 12-MAY-1994  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: Jackson Esq., David A.  
REGISTRATION NUMBER: 26,742  
REFERENCE/DOCKET NUMBER: 600-1-078  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 201 487-5800  
TELEFAX: 201 343-1684  
TELEX: 133521  
INFORMATION FOR SEQ ID NO: 25:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 94 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
HYPOTHETICAL: NO  
ANTI-SENSE: NO  
US-08-241-853-25

Query Match 23.4%; Score 35.5; DB 1; Length 94;  
Best Local Similarity 39.1%; Pred. No. 3.2e+02;  
Matches 9; Conservative 5; Mismatches 8; Indels 1; Gaps 1;

QY 4 LNLVAMNIIISKEKKEIKWIGLP 26

Db 5 LNTVSLN-ESSSFLEIRWVKPP 26

RESULT 40  
US-08-850-917-25  
Sequence 25, Application US/08850917  
Patent No. 5854045  
GENERAL INFORMATION:  
APPLICANT: Fang, Kathy S.  
APPLICANT: Hanafusa, Hidesaburo  
TITLE OF INVENTION: A TRANSMEMBRANE TYROSINE PHOSPHATASE  
NUMBER OF SEQUENCES: 37  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Klauber & Jackson  
STREET: 411 Hackensack Avenue  
CITY: Hackensack  
STATE: New Jersey  
COUNTRY: USA  
ZIP: 07601  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent In Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/850,917  
FILING DATE: 02-MAY-1997  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/241,853  
FILING DATE: 12-MAY-1994  
ATTORNEY/AGENT INFORMATION:  
NAME: Jackson Esq., David A.  
REGISTRATION NUMBER: 26,742  
REFERENCE/DOCKET NUMBER: 600-1-078  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 201 487-5800  
TELEFAX: 201 343-1684  
TELEX: 133521  
INFORMATION FOR SEQ ID NO: 25:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 94 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
HYPOTHETICAL: NO  
ANTI-SENSE: NO  
US-08-850-917-25

Query Match 23.4%; Score 35.5; DB 2; Length 94;  
Best Local Similarity 39.1%; Pred. No. 3.2e+02;  
Matches 9; Conservative 5; Mismatches 8; Indels 1; Gaps 1;

QY 4 LNLVAMNIIISKEKKEIKWIGLP 26  
DB 5 LNTVSLN-ESSSFLEIRWVKPP 26

RESULT 41  
US-09-308-935-9  
Sequence 9, Application US/09308935  
Patent No. 6268334  
GENERAL INFORMATION:  
APPLICANT: Bandara, Lashantha R  
APPLICANT: La Thangue, Nicholas B  
TITLE OF INVENTION: Peptide antagonists of DP transcription factors  
FILE REFERENCE: 620-67  
CURRENT APPLICATION NUMBER: US/09/308,935  
CURRENT FILING DATE: 1999-05-27  
EARLIER APPLICATION NUMBER: PCT/GB97/03506

us-09-900-147-6-rai

Tue Feb 17 11:55:54 2004

```

; EARLIER FILING DATE: 1997-12-22
; EARLIER APPLICATION NUMBER: GB 9626589.7
; EARLIER FILING DATE: 1996-12-20
; NUMBER OF SEQ ID NOS: 18
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 9
; LENGTH: 11
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Synthetic
US-09-308-935-9
Query Match      23.0%; Score 35; DB 3; Length 11;
Best Local Similarity 100.0%; Pred. No. 32;
Matches      7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 YDALNVL 7
DB      5 YDALNVL 11

RESULT 42
US-07-948-357-3
; Sequence 3, Application US/07948357
; Patent No. 5547932
; GENERAL INFORMATION:
; APPLICANT: Curiel, David T.
; APPLICANT: Birnstiel, Max L.
; APPLICANT: Cotten, Matthew
; APPLICANT: Wagner, Ernst
; APPLICANT: Zatloukal, Kurt
; APPLICANT: Plank, Christian
; APPLICANT: Oberhauser, Berndt
; APPLICANT: Schmidt, Walter G.M.
; TITLE OF INVENTION: Composition for Introducing Nucleic Acid
; TITLE OF INVENTION: Complexes Into Higher Eucaryotic Cells
; NUMBER OF SEQUENCES: 13
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Sterne, Kessler, Goldstein & Fox
; STREET: 1225 Connecticut Avenue
; CITY: Washington
; STATE: D.C.
; COUNTRY: U.S.A.
; ZIP: 20036
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/450,417
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/07/948,357
; FILING DATE: 19920923
; ATTORNEY/AGENT INFORMATION:
; NAME: Esmond, Robert W.
; REGISTRATION NUMBER: 32,893
; REFERENCE/DOCKET NUMBER: 0652.0940004
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 466-0800
; TELEFAX: (202) 833-8716
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 26 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: both
; MOLECULE TYPE: peptide
US-08-450-417-3
Query Match      23.0%; Score 35; DB 2; Length 26;
Best Local Similarity 50.0%; Pred. No. 85;
Matches      8; Conservative 2; Mismatches 6; Indels 0; Gaps 0;

QY      8 MAMNIISKEKKEIKWI 23
DB      1 MAQDIISTIGDLVRKI 16

RESULT 44
US-08-449-741-3
; Sequence 3, Application US/08449741
; Patent No. 6022735
; GENERAL INFORMATION:
; APPLICANT: Curiel, David T.
; APPLICANT: Birnstiel, Max L.
US-07-948-357-3
Query Match      23.0%; Score 35; DB 1; Length 26;
Best Local Similarity 50.0%; Pred. No. 85;
Matches      8; Conservative 2; Mismatches 6; Indels 0; Gaps 0;

QY      8 MAMNIISKEKKEIKWI 23
DB      1 MAQDIISTIGDLVRKI 16
```

APPLICANT: Cotten, Matthew  
APPLICANT: Wagner, Ernst  
APPLICANT: Zatloukal, Kurt  
APPLICANT: Plank, Christian  
APPLICANT: Oberhauser, Berndt  
APPLICANT: Schmidt, Walter G.M.  
TITLE OF INVENTION: Composition for Introducing Nucleic Acid  
NUMBER OF SEQUENCES: 13  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Sterne, Kessler, Goldstein & Fox  
STREET: 1100 New York Ave., Suite 600  
CITY: Washington  
STATE: D.C.  
COUNTRY: U.S.A.  
ZIP: 20005  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/449,741  
FILING DATE: 25-MAY-1995  
CLASSIFICATION: 514  
ATTORNEY/AGENT INFORMATION:  
NAME: Fleisher, Raz E.  
REGISTRATION NUMBER: 34,331  
REFERENCE/DOCKET NUMBER: 0652.0940007  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (202) 371-2600  
TELEFAX: (202) 371-2540  
INFORMATION FOR SEQ ID NO: 3:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 26 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: both  
MOLECULE TYPE: peptide  
US-08-449-741-3

Query Match 23.0%; Score 35; DB 3; Length 26;  
Best Local Similarity 50.0%; Pred. No. 85;  
Matches 8; Conservative 2; Mismatches 6; Indels 0; Gaps 0;

Qy 8 MAMNIIISKEKKEIKWI 23  
||:||||:||||  
Db 1 MAQDIISTIGDLVKWI 16

## RESULT 45

US-08-782-997A-5  
Sequence 5, Application US/08782997A  
Patent No. 603602  
GENERAL INFORMATION:  
APPLICANT: Legendre, Jean-Yves  
APPLICANT: Supersaxo, Andreas  
APPLICANT: Trzeciak, Arnold  
TITLE OF INVENTION: Peptide Conjugates for Transfecting  
NUMBER OF SEQUENCES: 37  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Hoffmann-La Roche Inc.  
STREET: 340 Kingdland Street  
CITY: Nutley  
STATE: New Jersey  
COUNTRY: U.S.A.  
ZIP: 07110  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25

CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/782,997A  
FILING DATE: 14-JAN-1997  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: EP 96100603.8  
FILING DATE: 17-JAN-1996  
ATTORNEY/AGENT INFORMATION:  
NAME: Kreisler, Lewis J  
REGISTRATION NUMBER: 38,522  
REFERENCE/DOCKET NUMBER: RAN 4600/73  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (973) 235-4387  
TELEFAX: (973) 235-2363  
INFORMATION FOR SEQ ID NO: 5:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 26 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
HYPOTHETICAL: NO  
FEATURE:  
NAME/KEY: Modified-site  
LOCATION: 1..2  
OTHER INFORMATION: /note= "Position 1 is fMet."  
US-08-782-997A-5

Query Match 23.0%; Score 35; DB 3; Length 26;  
Best Local Similarity 50.0%; Pred. No. 85;  
Matches 8; Conservative 2; Mismatches 6; Indels 0; Gaps 0;

Qy 8 MAMNIIISKEKKEIKWI 23  
||:||||:||||  
Db 1 MAQDIISTIGDLVKWI 16

Search completed: February 17, 2004, 10:59:41  
Job time : 21.6832 secs

THIS PAGE BLANK (user)

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: February 17, 2004, 10:50:12 ; Search time 57.0297 Seconds  
(without alignments)  
83.497 Million cell updates/sec

Title: US-09-900-147-6

Perfect score: 152

Sequence: 1 YDALNVLMANNIISKKEIKWIGLPTNSA 30

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1107863 seqs, 158726573 residues

Total number of hits satisfying chosen parameters: 717921

Minimum DB seq length: 0

Maximum DB seq length: 100

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : A Geneseq\_19Jun03.\*

```

1: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1980.DAT.*
2: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1981.DAT.*
3: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1982.DAT.*
4: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1983.DAT.*
5: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1984.DAT.*
6: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1985.DAT.*
7: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1986.DAT.*
8: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1987.DAT.*
9: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1988.DAT.*
10: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1989.DAT.*
11: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1990.DAT.*
12: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1991.DAT.*
13: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1992.DAT.*
14: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1993.DAT.*
15: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1994.DAT.*
16: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1995.DAT.*
17: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1996.DAT.*
18: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1997.DAT.*
19: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1998.DAT.*
20: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1999.DAT.*
21: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA2000.DAT.*
22: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA2001.DAT.*
23: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA2002.DAT.*
24: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA2003.DAT.*

```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Match	Length	ID	Description
1	152	100.0	30	19 AAW30507	DP-1 transcription
2	152	100.0	37	19 AAW30501	DP-1 transcription
3	101	66.4	20	19 AAW30505	DP-1 transcription
4	99	65.1	56	22 AAB67766	Fragment from a wh
5	90	59.2	83	20 AAY32163	Soybean DP-1 prote
6	80	52.6	15	19 AAW30510	DP-1 transcription
7	76	50.0	35	23 AAU72601	DEF domain consens
8	72	47.4	19	19 AAW30504	DP-1 transcription
9	68	44.7	19	19 AAW30515	DP-1 transcription

10	67	44.1	16	19 AAW30506	DP-1 transcription
11	67	44.1	28	19 AAW57051	E2F activity inhib
12	67	44.1	28	19 AAW57055	E2F activity inhib
13	64	42.1	19	19 AAW30516	DP-1 transcription
14	60	39.5	19	19 AAW30517	DP-1 transcription
15	59	38.8	93	23 AAU72561	Arabidopsis cell c
16	55.5	36.5	85	23 AAU78095	Human DNA binding
17	55.5	36.5	85	24 ABB82986	Repressor protein
18	55	36.2	14	19 AAW30511	DP-1 transcription
19	52	34.2	29	22 ABB03433	Human musculoskele
20	52	34.2	29	24 ABU12727	Novel human muscu
21	50.5	33.2	76	23 AAU78096	Human DNA binding
22	50.5	33.2	76	24 ABB82987	Repressor protein
23	49	32.2	15	19 AAW57052	E2F activity inhib
24	46.5	30.6	96	22 ABB58365	Human liver peptid
25	46.5	30.6	96	22 ABB42955	Peptide #10461 enc
26	46.5	30.6	96	22 AAM63859	Human brain expres
27	46.5	30.6	96	22 AAM76872	Human bone marrow
28	46.5	30.6	96	22 AAM36777	Peptide #10814 enc
29	46.5	30.6	96	23 ABB45881	Human peptide enco
30	45.5	29.9	80	20 AAY32167	Soybean E2F protei
31	43.5	28.6	28	21 AAY44493	Wheat E2F derived
32	43	28.3	51	23 ABB29212	Streptococcus poly
33	42	27.6	9	19 AAW30502	DP-1 transcription
34	42	27.6	57	20 AAY36705	Fragment of human
35	42	27.6	62	22 ABB17294	Human nervous syst
36	42	27.6	91	20 AAY37764	Amino acid sequenc
37	41.5	27.3	75	22 ABB49671	Human liver peptid
38	41.5	27.3	75	22 ABB29659	Peptide #2310 enco
39	41.5	27.3	75	22 ABB34837	Peptide #2343 enco
40	41.5	27.3	75	22 ABB20257	Protein #2256 enco
41	41.5	27.3	75	22 AAM55645	Human brain expres
42	41.5	27.3	75	22 AAM68025	Human bone marrow
43	41.5	27.3	75	22 AAM15843	Peptide #2277 enco
44	41.5	27.3	75	22 AAM28353	Peptide #2390 enco
45	41.5	27.3	75	22 AAM03581	Peptide #2263 enco

#### ALIGNMENTS

RESULT 1  
AAW30507  
ID AAW30507 standard; Peptide; 30 AA.  
XX AC AAW30507;  
XX DT 26-OCT-1998 (first entry)  
XX DP-1 transcription factor antagonist peptide H7.  
DE DP-1 transcription factor; antagonist; E2F protein; apoptosis;  
KW cell proliferation; cardiovascular cell; restenosis; tumour;  
KW surgical stent; therapy.  
XX OS Synthetic.  
OS Homo sapiens.  
XX  
FH Key Location/Qualifiers  
FT Peptide 3..9  
FT /note= "Claim 3"  
FT Peptide 5..15  
FT /note= "Claim 3"  
XX  
FN WO9828334-A1.  
XX  
PD 02-JUL-1998.  
XX  
PF 22-DEC-1997; 97WO-GB03506.  
XX  
PR 20-DEC-1996; 96GB-0026589.  
XX  
PA (PROL-) PROLIFIX LTD.

XX PI Bandara LR, La Thangue NB;  
 XX PT WPI; 1998-377596/32.  
 XX DR Polypeptide fragments of the DP-1 transcription factor - used for  
 XX PT inducing apoptosis, specifically in tumour and cardiovascular cells,  
 XX PT e.g. for preventing restenosis  
 XX PS Claim 4; Page 44; 55pp; English.  
 XX CC Peptide H7 comprises amino acid residues 170-199 in the DEF box  
 XX CC (I) (see AAW30501) of transcription factor DFI. Claimed peptides  
 XX CC (II) (see AAW30504-07) containing one or both of 2 motifs (see  
 XX CC AAW30502-03) of the DEF box are capable of antagonising the  
 XX CC heterodimerisation of a DP protein with an E2F protein. Also  
 XX CC claimed are variants of these peptides, especially containing  
 XX CC substitutions of residues corresponding to residues 167, 169, 171  
 XX CC and 175 of DP-1, fusion proteins (III) comprising (I) or (II) and a  
 XX CC membrane translocation sequence (see AAW30508), expression vectors  
 XX CC encoding (I)-(III) and host cells. (I)-(III) are used  
 XX CC therapeutically to induce apoptosis, specifically in tumour or  
 XX CC cardiovascular cells, either in vivo or in vitro, e.g. for purging  
 XX CC bone marrow. Surgical stents comprising (I)-(III) are used to  
 XX CC treat or prevent restenosis in patients who have undergone  
 XX CC angioplasty. (I)-(III) function by inactivating the DNA-binding  
 XX CC activity of DP/E2F heterodimers. They are also used as research  
 XX CC reagents, as positive controls in assays for identifying  
 XX CC antagonists of DP-1/E2F dimerisation and as immunoassay agents.  
 XX CC Also described is the use of sequences antisense to nucleic acids  
 XX CC encoding (I)-(III) to control DP levels in cells, particularly by  
 XX CC gene therapy. When formulated with cytotoxic or cytostatic agents,  
 XX CC (I)-(III) enhance cell killing.  
 XX SQ Sequence 30 AA;  
 Query Match 100.0%; Score 152; DB 19; Length 30;  
 Best Local Similarity 100.0%; Pred. No. 2.9e-17; Indels 0; Gaps 0;  
 Matches 30; Conservative 0; Mismatches 0;  
 QY 1 YDALNVLMMNNIISKEKEIKWIGLPTNSA 30  
 DB 1 YDALNVLMMNNIISKEKEIKWIGLPTNSA 30  
 RESULT 2  
 AAW30501  
 ID AAW30501 standard; Peptide; 37 AA.  
 XX AC AAW30501;  
 XX DT 26-OCT-1998 (first entry)  
 XX DE DP-1 transcription factor peptide H (DEF box).  
 XX KW DP-1; transcription factor; antagonist; E2F protein; apoptosis;  
 XX KW cell proliferation; cardiovascular cell; restenosis; tumour;  
 XX KW surgical stent; therapy.  
 XX OS Synthetic.  
 XX OS Homo sapiens.  
 XX PN WO9828334-A1.  
 XX PD 02-JUL-1998.  
 XX PF 22-DEC-1997; 97WO-GB03506.  
 XX PR 20-DEC-1996; 96GB-0026589.  
 XX PA (PROL-) PROLIFIX LTD.  
 XX PI Bandara LR, La Thangue NB;  
 XX DR WPI; 1998-377596/32.

XX DR WPI; 1998-377596/32.  
 XX PT Polypeptide fragments of the DP-1 transcription factor - used for  
 XX PT inducing apoptosis, specifically in tumour and cardiovascular cells,  
 XX PT e.g. for preventing restenosis  
 XX PS Claim 1; Page 44; 55pp; English.  
 XX CC Peptide H (I) comprises residues 163-199, i.e. the DEF box region,  
 XX CC of transcription factor DFI. Claimed fragments (II) (see AAW30502-07)  
 XX CC of (I) are capable of antagonising the heterodimerisation of a DP  
 XX CC protein with an E2F protein. Also claimed are fusion proteins  
 XX CC (III) comprising (I) or (II) and a membrane translocation sequence  
 XX CC (see AAW30508), expression vectors encoding (I)-(III) and host cells.  
 XX CC (I)-(III) are used therapeutically to induce apoptosis,  
 XX CC specifically in tumour or cardiovascular cells, either in vivo or in  
 XX CC vitro, e.g. for purging bone marrow. Surgical stents comprising  
 XX CC (I)-(III) are used to treat or prevent restenosis in patients who  
 XX CC have undergone angioplasty. (I)-(III) function by inactivating  
 XX CC the DNA-binding activity of DP/E2F heterodimers. They are also  
 XX CC used as research reagents, as positive controls in assays for  
 XX CC identifying antagonists of DP-1/E2F dimerisation and as immunoassay  
 XX CC agents. Also described is the use of sequences antisense to  
 XX CC nucleic acids encoding (I)-(III) to control DP levels in cells,  
 XX CC particularly by gene therapy. When formulated with cytotoxic  
 XX CC or cytostatic agents, (I)-(III) enhance cell killing.  
 XX SQ Sequence 37 AA;  
 Query Match 100.0%; Score 152; DB 19; Length 37;  
 Best Local Similarity 100.0%; Pred. No. 3.7e-17; Indels 0; Gaps 0;  
 Matches 30; Conservative 0; Mismatches 0;  
 QY 1 YDALNVLMMNNIISKEKEIKWIGLPTNSA 30  
 DB 8 YDALNVLMMNNIISKEKEIKWIGLPTNSA 37  
 RESULT 3  
 AAW30505  
 ID AAW30505 standard; Peptide; 20 AA.  
 XX AC AAW30505;  
 XX DT 26-OCT-1998 (first entry)  
 XX DE DP-1 transcription factor antagonist peptide H3.  
 XX KW DP-1; transcription factor; antagonist; E2F protein; apoptosis;  
 XX KW cell proliferation; cardiovascular cell; restenosis; tumour;  
 XX KW surgical stent; therapy.  
 XX OS Synthetic.  
 XX OS Homo sapiens.  
 XX PN WO9828334-A1.  
 XX PD 02-JUL-1998.  
 XX PF 22-DEC-1997; 97WO-GB03506.  
 XX PR 20-DEC-1996; 96GB-0026589.  
 XX PA (PROL-) PROLIFIX LTD.  
 XX PI Bandara LR, La Thangue NB;  
 XX DR WPI; 1998-377596/32.

XX Polypeptide fragments of the DP-1 transcription factor - used for  
PT inducing apoptosis, specifically in tumour and cardiovascular cells,  
PT e.g. for preventing restenosis  
XX  
XX Claim 4; Page 44; 55pp; English.  
XX  
CC Peptide H3 comprises amino acid residues 174-193 in the DEF box  
CC (I) (see AAW30501) of transcription factor Dp1. Claimed peptides  
CC (II) (see AAW30504-07) containing one or both of 2 motifs (see  
CC AAW30502-03) of the DEF box are capable of antagonising the  
CC heterodimerisation of a DP protein with an E2F protein. Also  
CC claimed are variants of these peptides, especially containing  
CC substitutions of residues corresponding to residues 167, 169, 171  
CC and 175 of DP-1, fusion proteins (III) comprising (I) or (II) and a  
CC membrane translocation sequence (see AAW30508), expression vectors  
CC encoding (I)-(III) and host cells. (I)-(III) are used  
CC therapeutically to induce apoptosis, specifically in tumour or  
CC cardiovascular cells, either in vivo or in vitro, e.g. for purging  
CC bone marrow. Surgical stents comprising (I)-(III) are used to  
CC treat or prevent restenosis in patients who have undergone  
CC angioplasty. (I)-(III) function by inactivating the DNA-binding  
CC activity of DP/E2F heterodimers. They are also used as research  
CC reagents, as positive controls in assays for identifying  
CC antagonists of DP-1/E2F dimerisation and as immunoassay agents.  
CC Also described is the use of sequences antisense to nucleic acids  
CC encoding (I)-(III) to control DP levels in cells, particularly by  
CC gene therapy. When formulated with cytotoxic or cytostatic agents,  
CC (I)-(III) enhance cell killing.  
XX  
XX Sequence 20 AA;  
XX  
XX Query Match 66.4%; Score 101; DB 19; Length 20;  
XX Best Local Similarity 100.0%; Pred. No. 3.1e-09;  
XX Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
XX  
QY 5 NVLMAMNIISKEKKEIKWIG 24  
Db |||||:|||||:|||||  
1 NVLMAMNIISKEKKEIKWIG 20  
XX  
XX RESULT 4  
XX AAB67766  
XX ID AAB67766 standard; peptide; 56 AA.  
XX  
XX AAB67766;  
XX  
XX 11-JUN-2001 (first entry)  
XX  
XX Fragment from a wheat E2F-dimerisation partner (DP) protein.  
XX  
XX E2F-dimerisation partner; DP protein; E2F transcription factor;  
XX G1 phase; S phase; cell cycle; retinoblastoma protein;  
XX alter cell proliferation.  
XX  
XX Triticum monococcum.  
XX  
XX WO200121644-A2.  
XX  
XX 29-MAR-2001.  
XX  
XX 25-SEP-2000; 2000WO-BP09325.  
XX  
XX 24-SEP-1999; 99ES-0002127.  
XX  
XX 11-NOV-1999; 99ES-0002474.  
XX  
XX (CNSJ ) CONSEJO SUPERIOR INVESTIGACIONES CIENTIF.  
XX  
XX Gutierrez-Armenta C, Ramirez-Parra E;  
XX  
XX WPI; 2001-257972/26.  
XX  
XX N-PSDB; AAF80148.  
XX

PT New isolated, enriched, cell free and/or recombinant nucleic acid  
PT useful for e.g. altering cell proliferation characteristic such as to  
PT alter plant cell, organ or tissue size -  
XX  
XX Claim 10; Page 50; 77pp; English.  
XX  
CC AAB67764-68 represent fragments of a wheat E2F-dimerisation partner  
CC (DP) protein. The protein acts as a plant E2F transcription factor.  
CC E2F and DP are two proteins that hetero-dimerise to form an active  
CC transcription factor that regulates G1 to S phase of the cell cycle,  
CC and later, the expression of genes required for S-phase progression.  
CC E2F and retinoblastoma protein also interact as a hetero-dimer in  
CC cells to suppress certain genes. This repression involves binding of  
CC the retinoblastoma protein to the E2F-DP dimer that is in turn bound  
CC to sites on DNA through the E2F DNA binding domain. DP proteins can  
CC be modulated to alter plant cell, organ or tissue shape, and  
CC particularly to alter cell proliferation characteristic such as to  
CC alter plant cell, organ or tissue size.  
XX  
XX Sequence 56 AA;  
XX  
XX Query Match 65.1%; Score 99; DB 22; Length 56;  
XX Best Local Similarity 68.0%; Pred. No. 2.3e-08;  
XX Matches 17; Conservative 6; Mismatches 2; Indels 0; Gaps 0;  
XX  
QY 1 YDALNVLAMNIISKEKKEIKWIGL 25  
Db |||||:|||||:|||||:|||||  
18 YDAFNVLIALRVIAKEKKEIRNMWGL 42  
XX  
XX RESULT 5  
XX AAY32163  
XX ID AAY32163 standard; Protein; 83 AA.  
XX  
XX AAY32163;  
XX  
XX 01-FEB-2000 (first entry)  
XX  
XX Soybean DP-1 protein fragment.  
XX  
XX DP-1; soybean; cell cycle regulatory protein;  
XX transcription factor; herbicide.  
XX  
XX Glycine max.  
XX  
XX Key Location/Qualifiers  
XX Misc-difference 10 /note= "encoded by GNC"  
XX Misc-difference 25 /note= "encoded by ANT"  
XX Misc-difference 26 /note= "encoded by GNC"  
XX Misc-difference 34 /note= "encoded by GGN"  
XX Misc-difference 35 /note= "encoded by CNT"  
XX Misc-difference 49 /note= "encoded by NAN"  
XX Misc-difference 54 /note= "encoded by ANG"  
XX Misc-difference 55 /note= "encoded by NAT"  
XX Misc-difference 59 /note= "encoded by NAG"  
XX Misc-difference 63 /note= "encoded by NAT"  
XX Misc-difference 69 /note= "encoded by GNG"  
XX Misc-difference 71 /note= "encoded by NTC"  
XX Misc-difference 80 /note= "encoded by CNA"  
XX

[illegible][illegible]



CC maize, tomato, alfalfa, oilseed rape, soybean, sunflower and canola.  
 CC CCP nucleic acid and polypeptide molecules are useful as modulating  
 CC agents in regulating cell cycle progression in plants. CCP is useful to  
 CC treat disorders characterised by insufficient or excessive production of  
 CC CCP protein or production of CCP protein forms which have decreased or  
 CC aberrant activity. Compounds that bind to or modulate the activity  
 CC of CCP polypeptide are useful as herbicides or plant growth regulators.  
 CC The polynucleotide is useful for modifying cell fate, plant development,  
 CC plant morphology, biochemistry and/or physiology, the length of the G1,  
 CC S, G2 and/or M phase of the cell cycle of a plant, initiation, promotion,  
 CC stimulation or enhancement of cell division, DNA replication, seed set,  
 CC seed size, seed development, tuber, fruit, leaf formation, shoot and root  
 CC initiation and/or development, nodule function, dwarfism in plants,  
 CC senescence, tolerance or resistance to stress. CCP, the polynucleotide  
 CC and the anti-CCP antibody are useful in agriculture to modulate the  
 CC protein levels or activity of a protein involved in the cell cycle due  
 CC to environmental conditions, including abiotic stress such as  
 CC cold, nutrient deprivation, heat, drought, salt stress, or biotic  
 CC stress such as pathogen attack, to modulate e.g. enhance crop yields,  
 CC and attenuate plant architecture, plant quality traits, plant  
 CC reproduction and seed development, endoreduplication in storage cells,  
 CC storage tissues and/or storage organs of plants or its parts. CCP is  
 CC useful as an immunogen to generate antibodies. CCP protein is useful to  
 CC screen for naturally occurring CCP substrates. The polynucleotide is  
 CC useful for expressing CCP protein, to detect CCP mRNA, or a genetic  
 CC lesion in a CCP gene and to modulate CCP activity. The present sequence  
 CC represents a motif which may be found in a CCP protein of the invention.  
 XX  
 XX Sequence 35 AA;

Query Match 50.0%; Score 76; DB 23; Length 35;  
 Best Local Similarity 64.0%; Pred. No. 6.7e-05;  
 Matches 16; Conservative 0; Mismatches 9; Indels 0; Gaps 0;

QY 2 DALNVLMAMNIIISKKEIKWIGLP 26  
 DB 11 DALNVXMAXXIXXXKKEIXWGLP 35

RESULT 8  
 AAW30504  
 ID AAW30504 standard; Peptide; 19 AA.  
 AC AAW30504;  
 XX  
 XX 26-OCT-1998 (first entry)  
 DE DP-1 transcription factor antagonist peptide H2.  
 XX  
 KW DP-1; transcription factor; antagonist; E2F protein; apoptosis;  
 KW cell proliferation; cardiovascular cell; restenosis; tumour;  
 KW surgical stent; therapy.  
 XX Synthetic.  
 OS Homo sapiens.  
 XX  
 XX Key Location/Qualifiers  
 FH Peptide 9..18  
 FT /note= "Claim 3"  
 FT  
 XX WO9828334-A1.  
 PN  
 XX 02-JUL-1998.  
 XX  
 XX 22-DEC-1997; 97WO-GB03506.  
 XX  
 XX 20-DEC-1996; 96GB-0026589.  
 XX  
 XX (PROL-) PROLIFIX LTD.  
 XX Bandara LR, La Thangue NB;  
 XX WPI; 1998-377596/32.  
 DR

XX Polypeptide fragments of the DP-1 transcription factor - used for  
 PT inducing apoptosis, specifically in tumour and cardiovascular cells,  
 PT e.g. for preventing restenosis  
 XX  
 PS Claim 4; Page 44; 55pp; English.  
 XX  
 CC Peptide H2 comprises amino acid residues 166-184 in the DEF box  
 CC (I) (see AAW30501) of transcription factor Dp1. Claimed peptides  
 CC (II) (see AAW30504-07) containing one or both of 2 motifs (see  
 CC AAW30502-03) of the DEF box are capable of antagonising the  
 CC heterodimerisation of a DP protein with an E2F protein. Also  
 CC claimed are variants of these peptides, especially containing  
 CC substitutions of residues corresponding to residues 167, 169, 171  
 CC and 175 of DP-1, fusion proteins (III) comprising (I) or (II) and a  
 CC membrane translocation sequence (see AAW30508), expression vectors  
 CC encoding (I)-(III) and host cells. (I)-(III) are used  
 CC therapeutically to induce apoptosis, specifically in tumour or  
 CC cardiovascular cells, either in vivo or in vitro, e.g. for purging  
 CC bone marrow. Surgical stents comprising (I)-(III) are used to  
 CC treat or prevent restenosis in patients who have undergone  
 CC angioplasty. (I)-(III) function by inactivating the DNA-binding  
 CC activity of DP/E2F heterodimers. They are also used as research  
 CC reagents, as positive controls in assays for identifying  
 CC antagonists of DP-1/E2F dimerisation and as immunoassay agents.  
 CC Also described is the use of sequences antisense to nucleic acids  
 CC encoding (I)-(III) to control DP levels in cells, particularly by  
 CC gene therapy. When formulated with cytotoxic or cytostatic agents,  
 CC (I)-(III) enhance cell killing.  
 XX

Sequence 19 AA;  
 Query Match 47.4%; Score 72; DB 19; Length 19;  
 Best Local Similarity 100.0%; Pred. No. 0.00014;  
 Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 YDALNVLMAMNIIISK 15  
 DB 5 YDALNVLMAMNIIISK 19

RESULT 9  
 AAW30515  
 ID AAW30515 standard; Peptide; 19 AA.  
 AC AAW30515;  
 XX  
 XX 26-OCT-1998 (first entry)  
 DE DP-1 transcription factor peptide H2mt1.  
 XX  
 KW DP-1; transcription factor; antagonist; E2F protein; apoptosis;  
 KW cell proliferation; cardiovascular cell; restenosis; tumour;  
 KW surgical stent; therapy.  
 XX Synthetic.  
 OS Homo sapiens.  
 XX  
 XX Key Location/Qualifiers  
 FH Misc-difference 4  
 FT /note= "V169A mutation"  
 FT Misc-difference 10  
 FT /note= "V175A mutation"  
 FT  
 XX WO9828334-A1.  
 PN  
 XX 02-JUL-1998.  
 XX  
 XX 22-DEC-1997; 97WO-GB03506.  
 XX  
 XX 20-DEC-1996; 96GB-0026589.  
 XX  
 XX (PROL-) PROLIFIX LTD.  
 PA

XX Bandara LR, La Thangue NB;  
 XX WPI; 1998-377596/32.  
 XX Polypeptide fragments of the DP-1 transcription factor - used for  
 PT inducing apoptosis, specifically in tumour and cardiovascular cells,  
 PT e.g. for preventing restenosis  
 XX Example D; Page 26; 55pp; English.  
 XX Peptide H2mt1 is based on peptide H2 (see AAW30504) from the DEF box  
 CC (see AAW30501) of transcription factor DP1. The H2mt1 peptide, in  
 CC which H2 residues corresponding to DP1 residues Val169 and Val175  
 CC are substituted by Ala residues, behaves in a similar fashion to  
 CC the wild-type H2 peptide in its ability to inactivate E2F site DNA  
 CC binding activity in D9 EC cell extracts. H2 is an antagonist of the  
 CC heterodimerisation of DP1 with E2F. Thus, the Val-169 and Val-175  
 CC residues of H2 play little role in this activity. H2 and other  
 CC claimed peptides (see AAW30504-07) from the DEF box region of DP1 can  
 CC be used to induce apoptosis, specifically in tumour and  
 CC cardiovascular cells, e.g. for the prevention of restenosis.  
 XX Sequence 19 AA;  
 SQ Query Match 44.7%; Score 68; DB 19; Length 19;  
 Best Local Similarity 93.3%; Pred. No. 0.00063;  
 Matches 14; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
 QY 1 YDALNVLMMNNIISK 15  
 DB 5 YDALNVLMMNNIISK 19  
 RESULT 10  
 AAW30506  
 ID AAW30506 standard; Peptide; 16 AA.  
 AC AAW30506;  
 XX 26-OCT-1998 (first entry)  
 XX DP-1 transcription factor antagonist peptide H5.  
 XX DP-1; transcription factor; antagonist; E2F protein; apoptosis;  
 KW cell proliferation; cardiovascular cell; restenosis; tumour;  
 KW surgical stent; therapy.  
 XX Synthetic.  
 OS Homo sapiens.  
 XX Key Location/Qualifiers  
 FH Peptide 5..11 /note= "Claim 3"  
 FT Peptide 7..16 /note= "Claim 3"  
 FT Peptide 7..16 /note= "Claim 3"  
 XX WO9828334-A1.  
 PN 02-JUL-1998.  
 XX 22-DEC-1997; 97WO-GB03506.  
 XX 20-DEC-1996; 96GB-0026589.  
 XX (PROL-) PROLIFIX LTD.  
 XX Bandara LR, La Thangue NB;  
 XX WPI; 1998-377596/32.  
 XX Polypeptide fragments of the DP-1 transcription factor - used for  
 PT inducing apoptosis, specifically in tumour and cardiovascular cells,  
 PT

PT e.g. for preventing restenosis  
 XX Claim 4; Page 44; 55pp; English.  
 XX Peptide H5 comprises amino acid residues 168-183 in the DEF box  
 CC (I) (see AAW30501) of transcription factor DP1. Claimed peptides  
 CC (II) (see AAW30504-07) containing one or both of 2 motifs (see  
 CC AAW30502-03) of the DEF box are capable of antagonising the  
 CC heterodimerisation of a DP protein with an E2F protein. Also  
 CC claimed are variants of these peptides, especially containing  
 CC substitutions of residues corresponding to residues 167, 169, 171  
 CC and 175 of DP-1, fusion proteins (III) comprising (I) or (II) and a  
 CC membrane translocation sequence (see AAW30508), expression vectors  
 CC encoding (I)-(III) and host cells. (I)-(III) are used  
 CC therapeutically to induce apoptosis, specifically in tumour or  
 CC cardiovascular cells, either in vivo or in vitro, e.g. for purging  
 CC bone marrow. Surgical stents comprising (I)-(III) are used to  
 CC treat or prevent restenosis in patients who have undergone  
 CC angioplasty. (I)-(III) function by inactivating the DNA-binding  
 CC activity of DP/E2F heterodimers. They are also used as research  
 CC reagents, as positive controls in assays for identifying  
 CC antagonists of DP-1/E2F dimerisation and as immunoassay agents.  
 CC Also described is the use of sequences antisense to nucleic acids  
 CC encoding (I)-(III) to control DP levels in cells, particularly by  
 CC gene therapy. When formulated with cytotoxic or cytostatic agents,  
 CC (I)-(III) enhance cell killing.  
 XX Sequence 16 AA;  
 SQ Query Match 44.1%; Score 67; DB 19; Length 16;  
 Best Local Similarity 100.0%; Pred. No. 0.00075;  
 Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 YDALNVLMMNNIIS 14  
 DB 3 YDALNVLMMNNIIS 16  
 RESULT 11  
 AAW57051  
 ID AAW57051 standard; peptide; 28 AA.  
 XX AAW57051;  
 XX 28-AUG-1998 (first entry)  
 XX E2F activity inhibiting compound Ib-1.  
 DE E2F activity; inhibitor; treatment; tumour; arteriosclerosis.  
 KW Synthetic.  
 OS Homo sapiens.  
 XX Key Location/Qualifiers  
 FH Modified-site 1 /note= "N-terminal acetyl"  
 FT Modified-site 28 /note= "C-terminal amide"  
 FT WO9814474-A1.  
 PN 09-APR-1998.  
 XX 26-SEP-1997; 97WO-JP03442.  
 XX 30-SEP-1996; 96JP-0259432.  
 XX (KYOW) KYOWA HAKKO KOGYO KK.  
 XX Mizukami T, Shibata K, Yamasaki M, Yoshida T;  
 XX WPI; 1998-240020/21.  
 XX

PT E2F activity inhibitors - for treatment and prevention of tumours  
 PT and arteriosclerosis

XX Example 3; Page 27; 52pp; Japanese.

XX This represents a compound that can inhibit E2F activity. The compound  
 CC is of the formula R1 - A - R2 where R1 is an optionally substituted  
 CC alkanoyl, allyl, hetero-arylcabonyl, alkoxycabonyl, aryloxyabonyl,  
 CC hetero-aryloxyabonyl, or H, R2 is OH, or optionally substituted alkoxy  
 CC or amino, and A is an E2F family dimer forming region or DNA binding  
 CC region, of at least 12 consecutive amino acids. Compounds of this formula  
 CC can be used to inhibit E2F activity, and are useful in the treatment and  
 CC prevention of tumours and arteriosclerosis.

XX Sequence 28 AA;  
 SQ Query Match 44.1%; Score 67; DB 19; Length 28;  
 Best Local Similarity 100.0%; Pred. No. 0.0015;  
 Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 YDALNVLMAMNIIS 14  
 Db 15 YDALNVLMAMNIIS 28  
 |||||

# RESULT 12

AAW57055  
 ID AAW57055 standard; peptide; 28 AA.

AC AAW57055;

XX 28-AUG-1998 (first entry)

XX E2F activity inhibiting compound IB-3.

XX E2F activity; inhibitor; treatment; tumour; arteriosclerosis.

OS Synthetic.  
 OS Homo sapiens.

FX Key Location/Qualifiers

FT Modified-site 1 /note= "N-terminal lauroyl"

FT Modified-site 28

FT /note= "C-terminal amide"

XX WO9814474-A1.

XX 09-APR-1998.

XX 26-SRP-1997; 97WO-JP03442.

XX 30-SEP-1996; 96JP-0259432.

XX (KYOW ) KYOWA HAKKO KOGYO KK.

XX Mizukami T, Shibata K, Yamasaki M, Yoshida T;  
 XX WPI; 1998-240020/21.

XX E2F activity inhibitors - for treatment and prevention of tumours  
 PT and arteriosclerosis

XX Example 7; Page 33; 52pp; Japanese.

XX This represents a compound that can inhibit E2F activity. The compound  
 CC is of the formula R1 - A - R2 where R1 is an optionally substituted  
 CC alkanoyl, allyl, hetero-arylcabonyl, alkoxycabonyl, aryloxyabonyl,  
 CC hetero-aryloxyabonyl, or H, R2 is OH, or optionally substituted alkoxy  
 CC or amino, and A is an E2F family dimer forming region or DNA binding  
 CC region, of at least 12 consecutive amino acids. Compounds of this formula  
 CC can be used to inhibit E2F activity, and are useful in the treatment and  
 CC prevention of tumours and arteriosclerosis.

XX Sequence 28 AA;  
 SQ Query Match 44.1%; Score 67; DB 19; Length 28;  
 Best Local Similarity 100.0%; Pred. No. 0.0015;  
 Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 YDALNVLMAMNIIS 14  
 Db 15 YDALNVLMAMNIIS 28  
 |||||

# RESULT 13

AAW30516  
 ID AAW30516 standard; Peptide; 19 AA.

AC AAW30516;

XX 26-OCT-1998 (first entry)

XX DP-1 transcription factor antagonist peptide H2mt2.

XX DP-1; transcription factor; E2F protein; apoptosis;  
 KW cell proliferation; cardiovascular cell; restenosis; tumour;  
 KW surgical stent; therapy.

XX Synthetic.

OS Homo sapiens.

XX Key Location/Qualifiers

FT Misc-difference 2 /note= "R167A mutation"

FT Misc-difference 6 /note= "D171A mutation"

XX WO9828334-A1.

XX 02-JUL-1998.

XX 22-DEC-1997; 97WO-GB03506.

XX 20-DEC-1996; 96GB-0026589.

XX (PROL-) PROLIFIX LTD.

XX Bandara LR, La Thangue NB;

XX WPI; 1998-377596/32.

XX Polyptide fragments of the DP-1 transcription factor - used for  
 PT inducing apoptosis, specifically in tumour and cardiovascular cells,  
 PT e.g. for preventing restenosis

XX Example D; Page 26; 55pp; English.

XX Peptide H2mt2 is based on peptide H2 (see AAW30504) from the DEF box  
 CC (see AAW30501) of transcription factor DP1. In H2mt2, amino acid  
 CC residues of H2 that correspond to DP1 residues Arg167 and Asp171  
 CC are substituted by Ala residues. H2 is an antagonist of the  
 CC heterodimerisation of DP1 with E2F. H2mt2 retains some, but not  
 CC all, of this antagonistic activity. H2 and other claimed peptides  
 CC (see AAW30504-07) from the DEF box region of DP1 can be used to  
 CC induce apoptosis, specifically in tumour and cardiovascular cells,  
 CC e.g. for the prevention of restenosis.

XX Sequence 19 AA;  
 SQ Query Match 42.1%; Score 64; DB 19; Length 19;  
 Best Local Similarity 93.3%; Pred. No. 0.0028;  
 Matches 14; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 YDALNVLMAMNIISK 15  
 |||||

Tue Feb 17 11:55:54 2004

us-09-900-147-6.rag

Db	5	YAALNVLAMNIISK 19	
RESULT 14			
AAW30517			
ID	AAW30517	standard; Peptide; 19 AA.	
XX			
AC	AAW30517;		
XX			
DT	26-OCT-1998	(first entry)	
XX			
DE	DP-1	transcription factor peptide H2mt3.	
XX			
KW	DP-1;	transcription factor; antagonist; E2F protein; apoptosis;	
KW	cell proliferation;	cardiovascular cell; restenosis; tumour;	
KW	surgical stent;	therapy.	
XX			
OS	Synthetic.		
OS	Homo sapiens.		
XX			
XX	Key	Location/Qualifiers	
FT	Misc-difference 8	/note= "L173R mutation"	
FT	Misc-difference 11	/note= "L176R mutation"	
FT			
XX	WO9828334-A1.		
PN			
XX			
PD	02-JUL-1998.		
XX			
PF	22-DEC-1997;	97WO-GB03506.	
XX			
PR	20-DEC-1996;	96GB-0026589.	
XX			
PA	(PROL-) PROLIFIX LTD.		
XX			
PI	Bandara LR,	La Thangue NB;	
XX			
DR	WPI;	1998-377596/32.	
XX			
XX	Polypeptide fragments of the DP-1	transcription factor - used for	
PT	inducing apoptosis,	specifically in tumour and cardiovascular cells,	
PT	e.g. for preventing restenosis		
XX			
PS	Example D;	Page 26; 55pp; English.	
XX			
CC	Peptide H2mt3 is based on peptide H2 (see AAW30504) from the DEF box		
CC	(see AAW30501) of transcription factor DP1. In H2mt3, amino acid		
CC	residues of H2 that correspond to DPL residues Leu173 and Leu176		
CC	are substituted by Arg residues. H2 is an antagonist of the		
CC	heterodimerisation of DP1 with E2F. H2mt3 has none of the		
CC	antagonistic activity of H2. H2 and other claimed peptides (see		
CC	AAW30504-07) from the DEF box region of DP1 can be used to induce		
CC	apoptosis, specifically in tumour and cardiovascular cells, e.g.		
CC	for the prevention of restenosis.		
XX			
SQ	Sequence	19 AA;	
	Query Match	39.5%; Score 60; DB 19; Length 19;	
	Best Local Similarity	86.7%; Pred. No. 0.012;	
	Matches	13; Conservative 0; Mismatches 2; Indels 0; Gaps 0;	
QY	1	YDALNVLAMNIISK 15	
	5	YDARNVRMAMNIISK 19	
Db			
RESULT 15			
AAU72561			
ID	AAU72561	standard; Protein; 93 AA.	
XX			
AC	AAU72561;		
XX			

DT	26-FEB-2002	(first entry)	
XX			
DE	Arabidopsis cell cycle protein AtDpa 121-293.		
XX			
KW	Cell cycle protein; CCP; cell cycle regulation; herbicide;		
KW	plant growth regulator; plant development; abiotic stress; biotic stress;		
KW	nutrient deprivation; pathogen attack; crop yield; immunogen; mutant;		
KW	mutein.		
XX			
OS	Arabidopsis thaliana.		
OS	Synthetic.		
XX			
PN	WO200185946-A2.		
XX			
PD	15-NOV-2001.		
XX			
PF	14-MAY-2001;	2001WO-IB01307.	
XX			
PR	12-MAY-2000;	2000US-204045P.	
XX			
PA	(CROP-) CROPDESIGN NV.		
XX			
PI	Inze D, Boudolf V, De Veylder L, Acosta JAT, Magyar Z;		
XX			
DR	WPI;	2002-062249/08.	
XX			
XX	New cell cycle protein and nucleic acid molecule encoding it useful for		
PT	regulating cell cycle progression in plants and for identifying		
PT	modulators which are useful as herbicides or plant growth regulators -		
XX			
PS	Example 10; Page 285; 316pp; English.		
XX			
CC	The invention relates to a novel cell cycle protein (CCP) and the		
CC	polynucleotides encoding them. CCP is useful for identifying a compound		
CC	which modulates the activity of the polypeptide and which binds to the		
CC	polypeptide and an anti-CCP antibody is useful for detecting the presence		
CC	of CCP in a sample. A CCP modulator is useful for modulating the cell		
CC	cycle or growth of a plant such as Arabidopsis thaliana, rice, wheat,		
CC	maize, tomato, alfalfa, oilseed rape, soybean, sunflower and canola.		
CC	CCP nucleic acid and polypeptide molecules are useful as modulating		
CC	agents in regulating cell cycle progression in plants. CCP is useful to		
CC	treat disorders characterised by insufficient or excessive production of		
CC	CCP protein or production of CCP protein forms which have decreased or		
CC	aberrant activity. Compounds that bind to or modulate the activity		
CC	of CCP polypeptide are useful as herbicides or plant growth regulators.		
CC	The polynucleotide is useful for modifying cell fate, plant development,		
CC	plant morphology, biochemistry and/or physiology, the length of the G1,		
CC	S, G2 and/or M phase of the cell cycle of a plant, initiation, promotion,		
CC	stimulation or enhancement of cell division, DNA replication, seed set,		
CC	seed size, seed development, tuber, fruit, leaf formation, shoot and root		
CC	initiation and/or development, nodule function, dwarfism in plants,		
CC	senescence, tolerance or resistance to stress. CCP, the polynucleotide		
CC	and the anti-CCP antibody are useful in agriculture to modulate the		
CC	protein levels or activity of a protein involved in the cell cycle due		
CC	to environmental conditions, including abiotic stress such as		
CC	cold, nutrient deprivation, heat, drought, salt stress, or biotic		
CC	stress such as pathogen attack, to modulate e.g. enhance crop yields,		
CC	and attenuate plant architecture, plant quality traits, plant		
CC	reproduction and seed development, endoreplication in storage cells,		
CC	storage tissues and/or storage organs of plants or its parts. CCP is		
CC	useful as an immunogen to generate antibodies. The polynucleotide is		
CC	screen for naturally occurring CCP substrates. The polynucleotide is		
CC	useful for expressing CCP protein, to detect CCP mRNA, or a genetic		
CC	lesion in a CCP gene and to modulate CCP activity. The present sequence		
CC	represents a deletion mutant of a CCP protein of the invention.		
XX			
SQ	Sequence	93 AA;	
	Query Match	38.8%; Score 59; DB 23; Length 93;	
	Best Local Similarity	66.7%; Pred. No. 0.12;	
	Matches	10; Conservative 4; Mismatches 1; Indels 0; Gaps 0;	
QY	12	IISKEKEIKWIGLP 26	



Tue Feb 17 11:55:54 2004

us-09-900-147-6.rag

```

ID AAW30511 standard; Peptide; 14 AA.
XX
AC AAW30511;
XX
XX DT 26-OCT-1998 (first entry)
XX
XX DE DP-1 transcription factor peptide H6.
XX
XX DP-1; transcription factor; antagonist; E2F protein; apoptosis;
XX cell proliferation; cardiovascular cell; restenosis; tumour;
XX surgical stent; therapy.
XX
XX OS Synthetic.
XX OS Homo sapiens.
XX
XX PN WO9828334-A1.
XX
XX PD 02-JUL-1998.
XX
XX PF 22-DEC-1997; 97WO-GB03506.
XX
XX PR 20-DEC-1996; 96GB-0026589.
XX
XX PA (PROL-) PROLIFIX LTD.
XX
XX PI Bandara LR, La Thangue NB;
XX
XX DR WPI; 1998-377596/32.
XX
XX CC Polypeptide fragments of the DP-1 transcription factor - used for
XX inducing apoptosis, specifically in tumour and cardiovascular cells,
XX e.g. for preventing restenosis
XX
XX PS Example C; Page 41; 55pp; English.
XX
XX CC Peptide H6 comprises amino acid residues 167-180 in the DEF box
XX region (see AAW30501) of transcription factor DP1. Unlike claimed
XX peptides (see AAW30504-07) that contain one or both of 2 motifs (see
XX AAW30502-03) of the DP1 DEF box, peptide H6 is not capable of
XX antagonising the heterodimerisation of a DP protein with an E2F
XX protein. The claimed peptides, their variants and fusion proteins
XX can be used to induce apoptosis, specifically in tumour and
XX cardiovascular cells, e.g. to prevent restenosis.
XX
XX SQ Sequence 14 AA;
XX
XX Query Match 36.2%; Score 55; DB 19; Length 14;
XX Best Local Similarity 100.0%; Pred. No. 0.056;
XX Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
XX QY 1 YDALNVLWAMN 11
XX
XX DB 4 YDALNVLWAMN 14
XX
XX RESULT 19
XX ABB03433
XX ID ABB03433 standard; Protein; 29 AA.
XX
XX AC ABB03433;
XX
XX DT 08-JAN-2002 (first entry)
XX
XX DE Human musculoskeletal system related polypeptide SEQ ID NO 1380.
XX
XX KW Cytostatic; immunosuppressive; nootropic; neuroprotective; antiviral;
XX antiallergic; hepatotropic; antidiabetic; antiinflammatory; antiulcer;
XX vulnary; anticonvulsant; antibacterial; antifungal; antiparasitic;
XX cardiac; gene therapy; cancer; immune disorder; cardiovascular disorder;
XX neurological disease; infection; human; secreted protein;
XX musculoskeletal system.
XX
XX OS Homo sapiens.
XX
```

```

XX
PN WO200155367-A1.
XX
XX PD 02-AUG-2001.
XX
XX PF 17-JAN-2001; 2001WO-US01338.
XX
XX PR 31-JAN-2000; 2000US-0179065.
XX PR 04-FEB-2000; 2000US-0180628.
XX PR 24-FEB-2000; 2000US-0184664.
XX PR 02-MAR-2000; 2000US-0186350.
XX PR 16-MAR-2000; 2000US-0189874.
XX PR 17-MAR-2000; 2000US-0190076.
XX PR 18-APR-2000; 2000US-0198123.
XX PR 19-MAY-2000; 2000US-0205515.
XX PR 28-JUN-2000; 2000US-0209467.
XX PR 30-JUN-2000; 2000US-0214886.
XX PR 07-JUL-2000; 2000US-0215135.
XX PR 07-JUL-2000; 2000US-0216647.
XX PR 11-JUL-2000; 2000US-0216880.
XX PR 11-JUL-2000; 2000US-0217487.
XX PR 14-JUL-2000; 2000US-0217496.
XX PR 26-JUL-2000; 2000US-0218290.
XX PR 26-JUL-2000; 2000US-0220963.
XX PR 14-AUG-2000; 2000US-0220964.
XX PR 14-AUG-2000; 2000US-0224518.
XX PR 14-AUG-2000; 2000US-0224519.
XX PR 14-AUG-2000; 2000US-0225213.
XX PR 14-AUG-2000; 2000US-0225214.
XX PR 14-AUG-2000; 2000US-0225266.
XX PR 14-AUG-2000; 2000US-0225267.
XX PR 14-AUG-2000; 2000US-0225268.
XX PR 14-AUG-2000; 2000US-0225270.
XX PR 14-AUG-2000; 2000US-0225447.
XX PR 14-AUG-2000; 2000US-0225757.
XX PR 14-AUG-2000; 2000US-0225758.
XX PR 14-AUG-2000; 2000US-0225759.
XX PR 18-AUG-2000; 2000US-0226279.
XX PR 22-AUG-2000; 2000US-0226681.
XX PR 22-AUG-2000; 2000US-0226868.
XX PR 23-AUG-2000; 2000US-0227182.
XX PR 30-AUG-2000; 2000US-0227009.
XX PR 01-SEP-2000; 2000US-0228924.
XX PR 01-SEP-2000; 2000US-0229287.
XX PR 01-SEP-2000; 2000US-0229343.
XX PR 01-SEP-2000; 2000US-0229344.
XX PR 05-SEP-2000; 2000US-0229345.
XX PR 05-SEP-2000; 2000US-0229509.
XX PR 05-SEP-2000; 2000US-0229513.
XX PR 06-SEP-2000; 2000US-0230437.
XX PR 06-SEP-2000; 2000US-0230438.
XX PR 08-SEP-2000; 2000US-0231242.
XX PR 08-SEP-2000; 2000US-0231243.
XX PR 08-SEP-2000; 2000US-0231244.
XX PR 08-SEP-2000; 2000US-0231413.
XX PR 08-SEP-2000; 2000US-0231414.
XX PR 08-SEP-2000; 2000US-0232080.
XX PR 08-SEP-2000; 2000US-0232081.
XX PR 12-SEP-2000; 2000US-0231968.
XX PR 14-SEP-2000; 2000US-0232397.
XX PR 14-SEP-2000; 2000US-0232398.
XX PR 14-SEP-2000; 2000US-0232399.
XX PR 14-SEP-2000; 2000US-0232400.
XX PR 14-SEP-2000; 2000US-0232401.
XX PR 14-SEP-2000; 2000US-0233063.
XX PR 14-SEP-2000; 2000US-0233064.
XX PR 14-SEP-2000; 2000US-0233065.
XX PR 21-SEP-2000; 2000US-0234223.
XX PR 21-SEP-2000; 2000US-0234274.
XX PR 25-SEP-2000; 2000US-0234997.
XX PR 25-SEP-2000; 2000US-0234998.
XX PR 26-SEP-2000; 2000US-0234984.
XX PR 27-SEP-2000; 2000US-0235834.
XX
```

PR 27-SEP-2000; 2000US-0235836.  
PR 29-SEP-2000; 2000US-0236327.  
PR 29-SEP-2000; 2000US-0236367.  
PR 29-SEP-2000; 2000US-0236368.  
PR 29-SEP-2000; 2000US-0236369.  
PR 29-SEP-2000; 2000US-0236370.  
PR 02-OCT-2000; 2000US-0236802.  
PR 02-OCT-2000; 2000US-0237037.  
PR 02-OCT-2000; 2000US-0237038.  
PR 02-OCT-2000; 2000US-0237039.  
PR 02-OCT-2000; 2000US-0237040.  
PR 13-OCT-2000; 2000US-0239935.  
PR 13-OCT-2000; 2000US-0239937.  
PR 20-OCT-2000; 2000US-0240960.  
PR 20-OCT-2000; 2000US-0241221.  
PR 20-OCT-2000; 2000US-0241785.  
PR 20-OCT-2000; 2000US-0241786.  
PR 20-OCT-2000; 2000US-0241787.  
PR 20-OCT-2000; 2000US-0241808.  
PR 20-OCT-2000; 2000US-0241809.  
PR 20-OCT-2000; 2000US-0241826.  
PR 01-NOV-2000; 2000US-0244617.  
PR 08-NOV-2000; 2000US-0246474.  
PR 08-NOV-2000; 2000US-0246475.  
PR 08-NOV-2000; 2000US-0246476.  
PR 08-NOV-2000; 2000US-0246477.  
PR 08-NOV-2000; 2000US-0246478.  
PR 08-NOV-2000; 2000US-0246523.  
PR 08-NOV-2000; 2000US-0246524.  
PR 08-NOV-2000; 2000US-0246525.  
PR 08-NOV-2000; 2000US-0246526.  
PR 08-NOV-2000; 2000US-0246527.  
PR 08-NOV-2000; 2000US-0246528.  
PR 08-NOV-2000; 2000US-0246532.  
PR 08-NOV-2000; 2000US-0246610.  
PR 08-NOV-2000; 2000US-0246611.  
PR 08-NOV-2000; 2000US-0246613.  
PR 17-NOV-2000; 2000US-0249207.  
PR 17-NOV-2000; 2000US-0249208.  
PR 17-NOV-2000; 2000US-0249209.  
PR 17-NOV-2000; 2000US-0249210.  
PR 17-NOV-2000; 2000US-0249211.  
PR 17-NOV-2000; 2000US-0249212.  
PR 17-NOV-2000; 2000US-0249213.  
PR 17-NOV-2000; 2000US-0249214.  
PR 17-NOV-2000; 2000US-0249215.  
PR 17-NOV-2000; 2000US-0249216.  
PR 17-NOV-2000; 2000US-0249217.  
PR 17-NOV-2000; 2000US-0249218.  
PR 17-NOV-2000; 2000US-0249244.  
PR 17-NOV-2000; 2000US-0249245.  
PR 17-NOV-2000; 2000US-0249246.  
PR 17-NOV-2000; 2000US-0249285.  
PR 17-NOV-2000; 2000US-0249287.  
PR 17-NOV-2000; 2000US-0249299.  
PR 17-NOV-2000; 2000US-0249300.  
PR 01-DEC-2000; 2000US-0250160.  
PR 01-DEC-2000; 2000US-0250391.  
PR 05-DEC-2000; 2000US-0251030.  
PR 05-DEC-2000; 2000US-0251988.  
PR 05-DEC-2000; 2000US-0256719.  
PR 06-DEC-2000; 2000US-0251479.  
PR 08-DEC-2000; 2000US-0251856.  
PR 08-DEC-2000; 2000US-0251868.  
PR 08-DEC-2000; 2000US-0251869.  
PR 08-DEC-2000; 2000US-0251989.  
PR 08-DEC-2000; 2000US-0251990.  
PR 11-DEC-2000; 2000US-0254097.  
PR 05-JAN-2001; 2001US-0259678.  
XX  
XX  
XX (HUMA-) HUMAN GENOME SCI INC.  
FA  
XX

PI Rosen CA, Barash SC, Ruben SM;  
XX WPI; 2001-451937/48.  
DR N-PSDB; AAL35015.  
XX  
XX Isolated polypeptide for treating, preventing and/ or prognosing  
PT disorders related to the musculoskeletal system including  
PT musculoskeletal cancers and also for testing and detection e.g.  
PT diagnosis -  
XX  
XX Claim 11; SEQ ID NO 1380; 781pp + Sequence Listing; English.  
XX  
CC The invention relates to novel genes (AAL34669-AAL37666) and proteins  
CC (AB03087-AB04109) associated with the musculoskeletal system useful  
CC for preventing, treating or ameliorating medical conditions e.g. by  
CC protein or gene therapy. The genes are isolated from a range of human  
CC tissues disclosed in the specification. The nucleic acids, proteins,  
CC antibodies and (ant)agonists are useful in the diagnosis, treatment  
CC and prevention of: (a) cancer, e.g. breast and ovarian cancer and  
CC other cancers of the adrenal gland, bone, bone marrow, breast,  
CC gastrointestinal tract, liver, lung, or urogenital; (b) immune  
CC disorders e.g. Addison's disease, allergies, autoimmune haemolytic  
CC anaemia, autoimmune thyroiditis, diabetes mellitus, Crohn's disease,  
CC multiple sclerosis, rheumatoid arthritis and ulcerative colitis;  
CC (c) cardiovascular disorders such as myocardial ischaemias; (d) wound  
CC healing; (e) neurological diseases e.g. cerebral anoxia and epilepsy;  
CC and (f) infectious diseases such as viral, bacterial, fungal and  
CC parasitic infections.  
CC Note: The sequence data for this patent did not form part of the  
CC printed specification, but was obtained in electronic format directly  
CC from WIPO at ftp.wipo.int/pub/published\_pct\_sequences.  
XX  
XX SQ Sequence 29 AA;  
Query Match 34.2%; Score 52; DB 22; Length 29;  
Best Local Similarity 62.5%; Pred. No. 0.41;  
Matches 10; Conservative 2; Mismatches 4; Indels 0; Gaps 0;  
Qy 10 MNIISKEKEIKWIGL 25  
| | | | | | | | | |  
Db 13 MKIFSKEKKIGWPGM 28  
| | | | | | | | | |  
RESULT 20  
ABU12727  
ID ABU12727 standard; Protein; 29 AA.  
XX  
XX ABU12727;  
AC  
XX  
XX DT 26-FEB-2003 (first entry)  
XX  
XX DE Novel human musculoskeletal system antigen #347.  
XX  
XX Musculoskeletal system antigen; cancer; metastasis;  
KW re-vascularisation; thrombosis; arteriosclerosis; mineral content;  
KW cardiovascular condition; wound; injury; burn; angiogenesis; ulcer;  
KW post-operative tissue repair; limb regeneration; neuronal growth;  
KW neurodegenerative disorder; Alzheimer's disease; Parkinson's disease;  
KW AIDS-related complex; chondrocyte growth; bone regeneration;  
KW periodontal regeneration; tissue transport; bone graft; skin aging;  
KW keratinocyte growth; hair loss; melanocyte growth; cell proliferation;  
KW cell growth; organ transplant; cell differentiation; body height;  
KW weight; hair colour; eye colour; skin; percentage of adipose tissue;  
KW pigmentation; cosmetic surgery; metabolism; biorhythm; cardiac rhythm;  
KW depression; tendency for violence; pain; reproductive capability;  
KW hormone level; endocrine level; appetite; libido; memory; stress;  
KW storage capability; fat content; lipid content; protein content;  
KW carbohydrate content; vitamin content; cofactor content;  
KW nutritional component.  
XX  
XX Homo sapiens.  
OS  
XX  
XX US2002147140-A1.  
FN

XX 10-OCT-2002.  
 PD 17-JAN-2001; 2001US-0764877.  
 PF 31-JAN-2000; 2000US-179065P.  
 XX 04-FEB-2000; 2000US-180628P.  
 PR 28-JUN-2000; 2000US-214886P.  
 PR 07-JUL-2000; 2000US-216647P.  
 PR 11-JUL-2000; 2000US-216880P.  
 PR 14-JUL-2000; 2000US-217487P.  
 PR 14-JUL-2000; 2000US-217496P.  
 PR 26-JUL-2000; 2000US-220963P.  
 PR 26-JUL-2000; 2000US-220964P.  
 PR 14-AUG-2000; 2000US-224518P.  
 PR 14-AUG-2000; 2000US-224519P.  
 PR 14-AUG-2000; 2000US-225267P.  
 PR 14-AUG-2000; 2000US-225268P.  
 PR 14-AUG-2000; 2000US-225270P.  
 PR 14-AUG-2000; 2000US-225447P.  
 PR 14-AUG-2000; 2000US-225757P.  
 PR 22-AUG-2000; 2000US-225758P.  
 PR 30-AUG-2000; 2000US-226868P.  
 PR 01-SEP-2000; 2000US-229287P.  
 PR 01-SEP-2000; 2000US-229343P.  
 PR 01-SEP-2000; 2000US-229344P.  
 PR 05-SEP-2000; 2000US-229345P.  
 PR 05-SEP-2000; 2000US-229509P.  
 PR 08-SEP-2000; 2000US-229513P.  
 PR 21-SEP-2000; 2000US-231413P.  
 PR 21-SEP-2000; 2000US-234223P.  
 PR 21-SEP-2000; 2000US-234274P.  
 PR 25-SEP-2000; 2000US-234997P.  
 PR 27-SEP-2000; 2000US-235834P.  
 PR 29-SEP-2000; 2000US-236327P.  
 PR 29-SEP-2000; 2000US-236367P.  
 PR 29-SEP-2000; 2000US-236368P.  
 PR 29-SEP-2000; 2000US-236369P.  
 PR 29-SEP-2000; 2000US-236370P.  
 PR 02-OCT-2000; 2000US-236802P.  
 PR 02-OCT-2000; 2000US-237037P.  
 PR 02-OCT-2000; 2000US-237038P.  
 PR 02-OCT-2000; 2000US-237039P.  
 PR 02-OCT-2000; 2000US-237040P.  
 PR 13-OCT-2000; 2000US-239935P.  
 PR 20-OCT-2000; 2000US-240960P.  
 PR 20-OCT-2000; 2000US-241785P.  
 PR 20-OCT-2000; 2000US-241809P.  
 PR 01-NOV-2000; 2000US-244617P.  
 PR 17-NOV-2000; 2000US-249299P.  
 PR 08-DEC-2000; 2000US-251856P.  
 PR 08-DEC-2000; 2000US-251868P.  
 PR 08-DEC-2000; 2000US-251869P.  
 XX (ROSE/) ROSEN C.A.  
 PA (RUBE/) RUBEN S.M.  
 PA (BARA/) BARASH S.C.  
 XX Rosen CA, Ruben SM, Barash SC;  
 PI WPI; 2003-128199/12.  
 XX N-PSDB; ABX58003.  
 DR Isolated nucleic acid molecules encoding musculoskeletal system  
 DR associated polypeptides, useful for detecting disorders, e.g. cancer -  
 XX Claim 11; SEQ ID NO 1380; 321pp; English.  
 PS The invention describes an isolated nucleic acid molecule comprising a  
 CC sequence encoding musculoskeletal system associated polypeptides useful  
 CC for detecting disorders, e.g., cancer or cancer metastases, in animals

CC or humans. The nucleic acid; stimulates re-vascularisation of ischaemic  
 CC tissues associated with conditions such as thrombosis, arteriosclerosis,  
 CC and other cardiovascular conditions; treats wounds due to injuries,  
 CC burns, post-operative tissue repair, and ulcers; stimulates angiogenesis  
 CC and limb regeneration; stimulates neuronal growth; can treat and prevent  
 CC neuronal damage occurring in certain disorders or neurodegenerative  
 CC conditions, such as, Alzheimer's disease, Parkinson's disease, and  
 CC AIDS-related complex; stimulates chondrocyte growth, thus they can be  
 CC used to enhance bone and periodontal regeneration and aid in tissue  
 CC transports or bone grafts; prevents skin aging due to sunburn by  
 CC stimulating keratinocyte growth; prevents hair loss, since FGF family  
 CC members activate hair-forming cells and promotes melanocyte growth;  
 CC stimulates growth and differentiation of hematopoietic cells and bone  
 CC marrow cells when used in combination with other cytokines; maintains  
 CC organs before transplantation or for supporting cell culture of primary  
 CC tissues; induces tissue of mesodermal origin to differentiate in early  
 CC embryos; increases or decreases the differentiation or proliferation of  
 CC embryonic stem cells, besides, haematopoietic lineage; modulates  
 CC mammalian characteristics, such as, body height, weight, hair colour, eye  
 CC colour, skin, percentage of adipose tissue, pigmentation, size, and shape  
 CC (e.g., cosmetic surgery); modulates mammalian metabolism; changes  
 CC mammal's mental state or physical state by influencing biorhythms,  
 CC cardiac rhythms, depression, tendency for violence, tolerance for pain,  
 CC reproductive capabilities, hormonal or endocrine levels, appetite,  
 CC libido, memory, or stress; increases or decreases storage capabilities,  
 CC fat content, lipid, protein, carbohydrate, vitamins, minerals, cofactors  
 CC or other nutritional components. This is the amino acid sequence of a  
 CC novel human musculoskeletal system antigen.  
 CC Note: The sequence data for this patent did not form part of the  
 CC printed specification, but was obtained in electronic format directly  
 CC from the US patent office at  
 CC ftp.seqdata.uspto.gov/sequence.html?DocID=20020147140.

XX Sequence 29 AA;

Query Match 34.2%; Score 52; DB 24; Length 29;  
 Best Local Similarity 62.5%; Pred. NO. 0.41;  
 Matches 10; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

QY 10 MNIISKEKEIKWIGL 25  
 DB 13 MKIFSKKKKIGWPGM 28

RESULT 21

AAU78096  
 ID AAU78096 standard; Protein; 76 AA.

XX AAU78096;

XX 18-JUN-2002 (first entry)

XX Human DNA binding domain E2F-4.

XX Human; telomerase reverse transcriptase; TERT; Site C; Progeria; burn;  
 KW repressor binding site; Hutchinson-Gilford syndrome; AIDS; cancer;  
 KW acquired immunodeficiency syndrome; cardiovascular disease; osteoporosis;  
 KW skin rejuvenation; immune senescence; bone marrow transplant; skin graft;  
 KW neoplastic disease; TERT minimal promoter; DNA binding domain; E2F-4.

XX Homo sapiens.

XX Key Location/Qualifiers  
 XX Domain 7..49  
 XX /label= DNA\_binding\_domain

XX WO200216657-A1.

XX 28-FEB-2002.

XX 17-AUG-2001; 2001WO-US25861.

XX 24-AUG-2000; 2000US-227865P.



PR 01-SEP-2000; 2000US-230174P.  
 XX 05-OCT-2000; 2000US-238345P.  
 PA (SIER-) SIERRA SCI INC.  
 XX Andrews WH, Foster CA, Fraser S, Mohammadpour H;  
 XX WPI; 2002-280952/32.  
 DR  
 XX  
 XX Modulating expression of telomerase reverse transcriptase (TERT) in a  
 PT cell, for regulating proliferative capacity of a cell, involves  
 PT modulating TERT transcription repression by Site C repressor binding  
 PT site -  
 XX  
 PS Disclosure; Page 8; 66pp; English.  
 XX  
 CC The present invention relates to a new method of modulating expression  
 CC of telomerase reverse transcriptase (TERT) from a TERT expression system  
 CC that includes a TERT promoter and a Site C repressor binding site. The  
 CC method of the invention involves modulating TERT transcription repression  
 CC by the Site C repressor binding site. The method of the invention is  
 CC useful for modulating expression of TERT for producing a mammalian  
 CC antibody. The method is also useful in a variety of different  
 CC applications, including immortalisation of cells, production of reagents  
 CC for use in life science research, therapeutic applications, and  
 CC therapeutic agent screening applications. Increasing TERT expression  
 CC delays natural telomeric shortening and/or increases telomeric length and  
 CC is useful for treating disease conditions such as Progeria or  
 CC Hutchinson-Gilford syndrome, acquired immunodeficiency syndrome (AIDS),  
 CC cardiovascular disease, osteoporosis, in skin rejuvenation and to inhibit  
 CC immune senescence. The method can be employed to lengthen telomeres of  
 CC osteoblast and osteoclast stem cells, encouraging bone replacement and  
 CC proper remodeling and reinforcement, and can thus be used in bone marrow  
 CC transplants for the treatment of cancer and skin grafts for burn  
 CC victims and as such the method improves the survival and effectiveness of  
 CC bone marrow and skin cell transplants. Decreasing TERT expression is  
 CC useful for treating cellular proliferative disease conditions, including  
 CC neoplastic disease conditions e.g. cancer. The present amino acid  
 CC sequence represents the human DNA binding domain E2F-4.  
 XX  
 XX Sequence 76 AA;  
 SQ  
 Query Match 33.2%; Score 50.5; DB 23; Length 76;  
 Best Local Similarity 42.3%; Pred. No. 2.3;  
 Matches 11; Conservative 5; Mismatches 9; Indels 1; Gaps 1;  
 QY 1 YDALNVLMMNNIISKE-KKEIKWIGL 25  
 DB 49 YDITNVLEGIGLEKKSNSIQWKG 74  
 RESULT 22  
 ABB82987  
 ID ABB82987 standard; Protein; 76 AA.  
 XX  
 AC ABB82987;  
 XX  
 DT 14-APR-2003 (first entry)  
 XX  
 DE Repressor protein E2F-4 DNA binding domain.  
 XX  
 XX Telomerase reverse transcriptase; TERT; Site C repressor; transcription;  
 KW cytosolic; immunostimulant; anti-HIV; vulnerary; telomerase; human;  
 KW repressor protein; E2F-4.  
 XX  
 OS Homo sapiens.  
 XX  
 PN WO2002101010-A2.  
 XX  
 PD 19-DEC-2002.  
 XX  
 PF 06-JUN-2002; 2002WO-US17959.  
 XX

PR 07-JUN-2001; 2001US-296992P.  
 XX (SIER-) SIERRA SCI INC.  
 PA  
 XX Fooster CA, Fraser S, Mohammadpour H, Andrews WH;  
 XX WPI; 2003-167401/16.  
 DR  
 XX  
 XX Modulating expression of telomerase reverse transcriptase TERT by  
 PT blocking repression of TERT transcripction, usefull for the diagnosis and  
 PT treatment of disorders associated with aberrant telomerase activity  
 PT such as cancer and HIV -  
 XX  
 PS Disclosure; Page 7; 47pp; English.  
 XX  
 CC The invention relates to modulating expression of telomerase reverse  
 CC transcriptase (TERT) expression system that includes a TERT promoter and  
 CC a Site C repressor binding site. The method involves modulating TERT  
 CC transcripction repression by the Site C repressor binding site. The  
 CC methods and compositions of the present invention are usefull for the  
 CC immortalization of cells, production of reagents in life science  
 CC research, therapeutic agent screening applications, diagnosis and  
 CC treatment of disorders associated with aberrant telomerase activity such  
 CC as cancer, Progeria, immune senescence, HIV, and in skin rejuvenation.  
 CC The present sequence represents the DNA binding domain of a repressor  
 CC protein E2F-4.  
 XX  
 SQ Sequence 76 AA;  
 Query Match 33.2%; Score 50.5; DB 24; Length 76;  
 Best Local Similarity 42.3%; Pred. No. 2.3;  
 Matches 11; Conservative 5; Mismatches 9; Indels 1; Gaps 1;  
 QY 1 YDALNVLMMNNIISKE-KKEIKWIGL 25  
 DB 49 YDITNVLEGIGLEKKSNSIQWKG 74  
 RESULT 23  
 AAW57052  
 ID AAW57052 standard; peptide; 15 AA.  
 XX  
 AC AAW57052;  
 XX  
 DT 28-AUG-1998 (first entry)  
 XX  
 DE E2F activity inhibiting compound Ib-2.  
 XX  
 KW E2F activity; inhibitor; treatment; tumour; arteriosclerosis.  
 XX  
 OS Synthetic.  
 OS Homo sapiens.  
 XX  
 FH Key Location/Qualifiers  
 FT Modified-site 1 /note= "N-terminal acetyl"  
 FT Modified-site 15 /note= "C-terminal amide"  
 FT  
 XX  
 PN WO9814474-A1.  
 XX  
 PD 09-APR-1998.  
 XX  
 PF 26-SEP-1997; 97WO-JP03442.  
 XX  
 PR 30-SEP-1996; 96JP-0259432.  
 XX  
 PA (KYOW ) KYOWA HAKKO KOGYO KK.  
 XX  
 PI Mizukami T, Shibata K, Yamasaki M, Yoshida T;  
 XX WPI; 1998-240020/21.  
 XX

PT E2F activity inhibitors - for treatment and prevention of tumours  
PT and arteriosclerosis  
XX  
PS Example 4; Page 28; 52pp; Japanese.  
XX  
CC This represents a compound that can inhibit E2F activity. The compound  
CC is of the formula R1 - A - R2 where R1 is an optionally substituted  
CC alkanyl, allyl, hetero-arylcabonyl, alkoxy carbonyl, aryloxy carbonyl,  
CC hetero-aryloxy carbonyl, or H, R2 is OH, or optionally substituted alkoxy  
CC or amino, and A is an E2F family dimer forming region or DNA binding  
CC region, of at least 12 consecutive amino acids. Compounds of this formula  
CC can be used to inhibit E2F activity, and are useful in the treatment and  
CC prevention of tumours and arteriosclerosis.  
XX  
XX  
SQ Sequence 15 AA;  
Query Match 32.2%; Score 49; DB 19; Length 15;  
Best Local Similarity 100.0%; Pred. No. 0.56; Mismatches 0; Indels 0; Gaps 0;  
Matches 10; Conservative 0;  
QY 1 YDALNVLMMAM 10  
DB 6 YDALNVLMMAM 15  
RESULT 24  
ABG58365  
ID ABG58365 standard; Peptide; 96 AA.  
XX  
AC ABG58365;  
XX  
XX  
DT 25-FEB-2003 (first entry)  
XX  
DE Human liver peptide, SEQ ID No 37013.  
XX  
KW Human; liver; cirrhosis; hyperlipoproteinaemia; hyperlipidaemia;  
KW hypercholesterolaemia; coronary heart disease.  
XX  
OS Homo sapiens.  
XX  
XX  
PN WO200157273-A2.  
XX  
PD 09-AUG-2001.  
XX  
PF 30-JAN-2001; 2001WO-US00664.  
XX  
PR 04-FEB-2000; 2000US-0180312.  
PR 26-MAY-2000; 2000US-0207456.  
PR 30-JUN-2000; 2000US-0608408.  
PR 03-AUG-2000; 2000US-0632366.  
PR 21-SEP-2000; 2000US-0234687.  
PR 27-SEP-2000; 2000US-0236359.  
PR 04-OCT-2000; 2000GB-0024263.  
XX  
PA (MOLE-) MOLECULAR DYNAMICS INC.  
XX  
PI Penn SG, Hanzel DK, Chen W, Rank DR;  
XX  
XX WPI; 2001-48898/53.  
XX  
PT Human genome-derived single exon nucleic acid probes useful for  
PT analysing gene expression in human adult liver -  
XX  
PS Claim 27; SEQ ID No 37013; 658pp; English.  
XX  
CC The invention relates to a single exon nucleic acid probe (SNP) (I) for  
CC measuring human gene expression in a sample derived from human adult  
CC liver, comprising one of 13109 defined nucleotide sequences given in the  
CC specification (or complements/ fragments). The probe hybridises at high  
CC stringency to a nucleic acid molecule expressed in the human adult  
CC liver. (I) may be used for predicting, measuring and displaying gene  
CC expression in samples derived from human adult liver. The genes  
CC identified may be involved in genetic liver diseases such as cirrhosis,

CC hyperlipoproteinaemia, hyperlipidaemia and hypercholesterolaemia which  
CC is associated with coronary heart disease. ABG47348-ABG59930 represent  
CC human liver single exon encoded peptides of the invention.  
CC Note: The sequence information for this patent does not appear in the  
CC printed specification but was obtained in electronic format directly  
CC from WIPO at ftp.wipo.int/pub/published\_pct\_sequences.  
XX  
XX Sequence 96 AA;  
SQ  
Query Match 30.6%; Score 46.5; DB 22; Length 96;  
Best Local Similarity 36.0%; Pred. No. 13;  
Matches 9; Conservative 8; Mismatches 7; Indels 1; Gaps 1;  
QY 1 YDALNVLMMAMTIISK-EKKEIKWIG 24  
DB 7 YDIVNVLESLLHLSRVAKNQYGMHG 31  
RESULT 25  
ABB42955  
ID ABB42955 standard; Peptide; 96 AA.  
XX  
AC ABB42955;  
XX  
DT 04-FEB-2002 (first entry)  
XX  
DE Peptide #10461 encoded by human foetal liver single exon probe.  
XX Human; foetal liver; gene expression; single exon nucleic acid probe.  
XX Homo sapiens.  
XX WO200157277-A2.  
XX 09-AUG-2001.  
XX 30-JAN-2001; 2001WO-US00669.  
XX 04-FEB-2000; 2000US-0180312.  
XX 26-MAY-2000; 2000US-0207456.  
XX 30-JUN-2000; 2000US-0608408.  
XX 03-AUG-2000; 2000US-0632366.  
XX 21-SEP-2000; 2000US-0234687.  
XX 27-SEP-2000; 2000US-0236359.  
XX 04-OCT-2000; 2000GB-0024263.  
XX (MOLE-) MOLECULAR DYNAMICS INC.  
XX Penn SG, Hanzel DK, Chen W, Rank DR;  
XX WPI; 2001-483447/52.  
XX Human genome-derived single exon nucleic acid probes useful for  
XX analysing gene expression in human foetal liver -  
XX  
XX Claim 27; SEQ ID NO 35590; 639pp + sequence listing; English.  
XX  
XX The invention relates to a single exon nucleic acid probe for  
XX measuring human gene expression in a sample derived from human foetal  
XX liver. The single exon nucleic acid probes may be used for predicting,  
XX measuring and displaying gene expression in samples derived from human  
XX foetal liver. The present sequence is a peptide encoded by a single exon  
XX nucleic acid probe of the invention.  
XX Note: The sequence data for this patent did not form part of the  
XX printed specification, but was obtained in electronic format directly  
XX from WIPO at ftp.wipo.int/pub/published\_pct\_sequences.  
XX  
XX Sequence 96 AA;  
SQ  
Query Match 30.6%; Score 46.5; DB 22; Length 96;  
Best Local Similarity 36.0%; Pred. No. 13;  
Matches 9; Conservative 8; Mismatches 7; Indels 1; Gaps 1;

```

DE Human bone marrow expressed probe encoded protein SEQ ID NO: 36978.
XX
XX
KW Human; bone marrow expressed exon; gene expression analysis; probe;
KW microarray; cancer; leukaemia; lymphoma; myeloma.
XX
OS Homo sapiens.
XX
XX WO200157276-A2.
XX
XX 09-AUG-2001.
XX
XX 30-JAN-2001; 2001WO-US00668.
XX
XX 04-FEB-2000; 2000US-0180312.
XX 26-MAY-2000; 2000US-0207456.
XX 30-JUN-2000; 2000US-0608408.
XX 03-AUG-2000; 2000US-0632366.
XX 21-SEP-2000; 2000US-0234587.
XX 27-SEP-2000; 2000US-0236359.
XX 04-OCT-2000; 2000GB-0024263.
XX
XX (MOLE-) MOLECULAR DYNAMICS INC.
XX
XX Penn SG, Hanzel DK, Chen W, Rank DR;
XX
XX WPI; 2001-488900/53.
XX
XX Human genome-derived single exon nucleic acid probes useful for
XX analyzing gene expression in human bone marrow -
XX
XX Example 4; SEQ ID NO: 36978; 658pp + Sequence Listing; English.
XX
XX The present invention provides a number of single exon nucleic acid
XX probes which are derived from genomic sequences expressed in the human
XX bone marrow. They can be used to measure gene expression in bone marrow
XX samples, which may enable the improved diagnosis and treatment of cancers
XX such as lymphoma, leukaemia and myeloma. The present sequence is a
XX protein encoded by one of the probes of the invention.
XX
XX Sequence 96 AA;
XX
XX Query Match 30.6%; Score 46.5; DB 22; Length 96;
XX Best Local Similarity 36.0%; Pred. No. 13;
XX Matches 9; Conservative 8; Mismatches 7; Indels 1; Gaps 1;
XX
XX QY 1 YDNLNVLMMNNIISK-EKKEIKWIG 24
XX |||:::|:::|:::|:::|:::|
XX DB 7 YDINVLDESILHSRVANQYGMHG 31
XX
XX
XX RESULT 28
XX AAM36777
XX ID AAM36777 standard; Protein; 96 AA.
XX
XX AC AAM36777;
XX
XX DT 17-OCT-2001 (first entry)
XX
XX DE Peptide #10814 encoded by probe for measuring placental gene expression.
XX
XX KW Probe; microarray; human; placenta; antenatal diagnosis;
XX KW genetic disorder.
XX
XX OS Homo sapiens.
XX
XX PN WO200157272-A2.
XX
XX 09-AUG-2001.
XX
XX 30-JAN-2001; 2001WO-US00663.
XX
XX 04-FEB-2000; 2000US-0180312.
XX 26-MAY-2000; 2000US-0207456.
XX

```

30-JUN-2000; 2000US-0608408.  
03-AUG-2000; 2000US-0632366.  
21-SEP-2000; 2000US-0234687.  
27-SEP-2000; 2000US-0236359.  
04-OCT-2000; 2000GB-0024263.  
(MOLE-) MOLECULAR DYNAMICS INC.  
Penn SG, Hanzel DK, Chen W, Rank DR;  
WPI; 2001-488897/53.  
Human genome-derived single exon nucleic acid probes useful for  
analyzing gene expression in human placenta -  
Claim 27; SEQ ID No 37046; 654pp; English.  
The present invention relates to single exon nucleic acid probes (SENP;  
see A131315-A157546). The present sequence is a peptide encoded by one  
such probe. The probes are useful for producing a microarray for  
predicting, measuring and displaying gene expression in samples derived  
from human placenta. The probes are useful for antenatal diagnosis of  
human genetic disorders.  
Sequence 96 AA;  
Query Match 30.6%; Score 46.5; DB 22; Length 96;  
Best Local Similarity 36.0%; Pred. No. 13;  
Matches 9; Conservative 8; Mismatches 7; Indels 1; Gaps 1;  
Qy 1 YDALNVLMMNIIISK-EKKEIKWIG 24  
Db 7 YDIVNVLESLHLSRVAKNQYGMHG 31  
RESULT 29  
ABG45881  
ID ABG45881 standard; Peptide; 96 AA.  
AC ABG45881;  
XX 19-AUG-2002 (first entry)  
DT Human peptide encoded by genome-derived single exon probe SEQ ID 35546.  
DE Human; single exon probe; asthma; lung cancer; COPD; ILD;  
KW chronic obstructive pulmonary disease; interstitial lung disease;  
KW familial idiopathic pulmonary fibrosis; neurofibromatosis;  
KW tuberous sclerosis; Gaucher's disease; Niemann-Pick disease;  
KW Hermansky-Pudlak syndrome; sarcoidosis; pulmonary haemorrhoidosis;  
KW pulmonary histiocytosis; lymphangioleiomyomatosis; Karagener syndrome;  
KW pulmonary alveolar proteinosis; fibrocystic pulmonary dysplasia;  
KW primary ciliary dyskinesia; pulmonary hypertension;  
KW hyaline membrane disease.  
OS Homo sapiens.  
XX WO200186003-A2.  
PN 15-NOV-2001.  
PD 30-JAN-2001; 2001WO-US00665.  
PF 04-FEB-2000; 2000US-180312P.  
PR 26-MAY-2000; 2000US-207456P.  
PR 30-JUN-2000; 2000US-0608408.  
PR 03-AUG-2000; 2000US-0632366.  
PR 21-SEP-2000; 2000US-234687P.  
PR 27-SEP-2000; 2000US-236359P.  
PR 04-OCT-2000; 2000GB-0024263.  
(MOLE-) MOLECULAR DYNAMICS INC.  
XX

Penn SG, Hanzel DK, Chen W, Rank DR;  
WPI; 2002-114183/15.  
Spatially-addressable set of single exon nucleic acid probes, used to  
measure gene expression in human lung samples -  
Claim 27; SEQ ID No 35546; 634pp; English.  
The invention relates to a spatially-addressable set of single exon  
nucleic acid probes for measuring gene expression in a sample derived  
from human lung comprising single exon nucleic acid probes having one of  
12614 nucleic acid sequences mentioned in the specification, or their  
complements or the 12387 open reading frames derived from the 12614  
probes. Also included are a microarray comprising the novel set of  
probes; the novel set of probes which hybridise at high stringency to a  
nucleic acid expressed in the human lung; measuring gene expression in a  
sample derived from human lung, comprising (a) contacting the array with  
a collection of detectably labeled nucleic acids derived from human lung  
mRNA, and (b) measuring the label detectably bound to each probe of  
the array; identifying exons in a eukaryotic genome, comprising  
(a) algorithmically predicting at least one exon from genomic sequences  
of the eukaryote; and (b) detecting specific hybridisation of detectably  
labeled nucleic acids from eukaryote lung mRNA, to a single exon probe,  
having a fragment identical to the predicted exon, the probe is included  
in the above mentioned microarray; assigning exons to a single gene,  
comprising (a) identifying exons from genomic sequence by the method  
above and (b) measuring the expression of each of the exons in several  
tissues and/or cell types using hybridisation to a single exon  
microarrays having a probe with the exon, where a common pattern of  
expression of the exons in the tissues and/or cell types indicates that  
the exons should be assigned to a single gene; a peptide comprising one  
of 12011 sequences, mentioned in the specification, or encoded by the  
probes/open reading frames (ORF). The probes are used for gene  
expression analysis, and for identifying exons in a gene, particularly  
using human lung derived mRNA and for the study of lung diseases  
such as asthma, lung cancer, chronic obstructive pulmonary disease  
(COPD), interstitial lung disease (ILD), familial idiopathic pulmonary  
fibrosis, neurofibromatosis, tuberous sclerosis, Gaucher's disease,  
Niemann-Pick disease, Hermansky-Pudlak syndrome, sarcoidosis, pulmonary  
haemorrhoidosis, pulmonary histiocytosis, lymphangioleiomyomatosis,  
pulmonary alveolar proteinosis, Karagener syndrome, fibrocystic  
pulmonary dysplasia, primary ciliary dyskinesia, pulmonary hypertension  
and hyaline membrane disease. The present sequence is a peptide/protein  
encoded by a single exon probe of the invention.  
Note: The sequence data for this patent did not form part  
of the printed specification, but was obtained in electronic  
format directly from WIPO at  
ftp.wipo.int/pub/published\_pct\_sequences.  
Sequence 96 AA;  
Query Match 30.6%; Score 46.5; DB 23; Length 96;  
Best Local Similarity 36.0%; Pred. No. 13;  
Matches 9; Conservative 8; Mismatches 7; Indels 1; Gaps 1;  
Qy 1 YDALNVLMMNIIISK-EKKEIKWIG 24  
Db 7 YDIVNVLESLHLSRVAKNQYGMHG 31  
RESULT 30  
AAY32167  
ID AAY32167 standard; Protein; 80 AA.  
XX  
AC AAY32167;  
XX  
DT 01-FEB-2000 (first entry)  
XX Soybean E2F protein fragment.  
DE  
XX E2F; soybean; cell cycle regulatory protein; transcription factor;  
KW herbicide.





```
PR 18-DEC-1997; 97US-0070923.
PR 19-DEC-1997; 97US-0068169.
PR 19-DEC-1997; 97US-0068365.
PR 19-DEC-1997; 97US-0068367.
PR 19-DEC-1997; 97US-0068368.
XX (HUMA-) HUMAN GENOME SCI INC.
XX
XX Carter KC, Duan RD, Feng P, Ferrie AM, Florence C;
PI Florence K, Greene JW, Janat F, Kyaw H, Moore PA;
PI Ni J, Rosen CA, Ruben SM, Shi Y, Soppet DR, Wei Y;
PI Yu G;
XX WPI; 1999-418749/35.
XX
XX New isolated human genes encoding secreted polypeptides
PT
PS Disclosure; Page 526; 537pp; English.
XX
XX AAX97916 to AAX98029 represent 110 isolated human secreted protein
CC genes. AAY36224 to AAY36727 represent the secreted proteins encoded by
CC the 110 human genes. The genes and their corresponding secreted
CC polypeptides are useful for preventing, treating or ameliorating medical
CC conditions, e.g. by protein or gene therapy. Also pathological conditions
CC can be diagnosed by determining the amount of the new polypeptides in a
CC sample or by determining the presence of mutations in the new genes.
CC Specific uses are described for each of the 110 genes, based on which
CC tissues they are most highly expressed in, and include developing
CC products for the diagnosis or treatment of cancer, tumours, developmental
CC abnormalities and foetal deficiencies, blood disorders, diseases of the
CC immune system, autoimmune diseases, inflammation, allergies, Alzheimer's
CC and cognitive disorders, schizophrenia, arthritis, asthma, psoriasis,
CC sepsis, skin disorders, atherosclerosis, diabetes, cardiovascular
CC disorders, kidney disorders, digestive/endocrine disorders, infections
CC and AIDS. The polypeptides are also useful for identifying their binding
CC partners. The sequences given in AAX97907 to AAX97915 and AAY36223 are
CC used in the amplification of the present invention.
XX
XX Sequence 57 AA;
SQ
Query Match 27.6%; Score 42; DB 20; Length 57;
Best Local Similarity 29.2%; Pred. No. 38;
Matches 7; Conservative 7; Mismatches 10; Indels 0; Gaps 0;
QY 5 NVLMAMNIIISKKEIKWIGLPTN 28
Db 1 NFCVSKNTFNVRKPIKWKVIFAN 24
RESULT 35
ABBI7294
ID ABB17294 standard; Protein; 62 AA.
XX
XX ABB17294;
AC
XX
XX 23-JAN-2002 (first entry)
DT
XX
XX Human nervous system related polypeptide SEQ ID NO 5951.
DE
XX
XX Human; neurotropic; neuroprotective; cytostatic; dermatological; virucide;
KW immunosuppressive; antiinflammatory; anti-HIV; antibacterial; vulnerary;
KW antiparkinsonian; antischlicking; antianemic; antiarthritic; cancer;
KW antirheumatic; hepatotropic; cerebroprotective; antiinflammatory;
KW antiallergic; antidiabetic; antitumor; anticonvulsant; antifungal;
KW antiparasitic; cardiac; immune disorder; cardiovascular disorder;
KW neurological disease; infection; nephrotropic; gene therapy; vaccine.
XX
XX Homo sapiens.
OS
XX
XX WO200159063-A2.
PN
XX
XX 16-AUG-2001.
PD
XX
```

```
PF 17-JAN-2001; 2001WO-US01334.
XX
XX 31-JAN-2000; 2000US-0179065.
PR 04-FEB-2000; 2000US-0180628.
PR 24-FEB-2000; 2000US-0184664.
PR 02-MAR-2000; 2000US-0186350.
PR 16-MAR-2000; 2000US-0189874.
PR 17-MAR-2000; 2000US-0190076.
PR 18-APR-2000; 2000US-0198123.
PR 19-MAY-2000; 2000US-0205515.
PR 07-JUN-2000; 2000US-0209467.
PR 28-JUN-2000; 2000US-0214886.
PR 30-JUN-2000; 2000US-0215135.
PR 07-JUL-2000; 2000US-0216647.
PR 07-JUL-2000; 2000US-0216880.
PR 11-JUL-2000; 2000US-0217487.
PR 11-JUL-2000; 2000US-0217496.
PR 14-JUL-2000; 2000US-0218290.
PR 26-JUL-2000; 2000US-0220963.
PR 26-JUL-2000; 2000US-0220964.
PR 14-AUG-2000; 2000US-0224518.
PR 14-AUG-2000; 2000US-0224519.
PR 14-AUG-2000; 2000US-0225213.
PR 14-AUG-2000; 2000US-0225214.
PR 14-AUG-2000; 2000US-0225266.
PR 14-AUG-2000; 2000US-0225267.
PR 14-AUG-2000; 2000US-0225268.
PR 14-AUG-2000; 2000US-0225270.
PR 14-AUG-2000; 2000US-0225447.
PR 14-AUG-2000; 2000US-0225757.
PR 14-AUG-2000; 2000US-0225758.
PR 14-AUG-2000; 2000US-0225759.
PR 18-AUG-2000; 2000US-0226279.
PR 22-AUG-2000; 2000US-0226681.
PR 22-AUG-2000; 2000US-0226868.
PR 22-AUG-2000; 2000US-0227182.
PR 23-AUG-2000; 2000US-0227009.
PR 30-AUG-2000; 2000US-0228924.
PR 01-SEP-2000; 2000US-0229287.
PR 01-SEP-2000; 2000US-0229343.
PR 01-SEP-2000; 2000US-0229344.
PR 01-SEP-2000; 2000US-0229345.
PR 05-SEP-2000; 2000US-0229509.
PR 05-SEP-2000; 2000US-0229513.
PR 06-SEP-2000; 2000US-0230437.
PR 06-SEP-2000; 2000US-0230438.
PR 08-SEP-2000; 2000US-0231242.
PR 08-SEP-2000; 2000US-0231243.
PR 08-SEP-2000; 2000US-0231244.
PR 08-SEP-2000; 2000US-0231413.
PR 08-SEP-2000; 2000US-0231414.
PR 08-SEP-2000; 2000US-0232080.
PR 08-SEP-2000; 2000US-0232081.
PR 12-SEP-2000; 2000US-0231968.
PR 14-SEP-2000; 2000US-0232397.
PR 14-SEP-2000; 2000US-0232398.
PR 14-SEP-2000; 2000US-0232399.
PR 14-SEP-2000; 2000US-0232400.
PR 14-SEP-2000; 2000US-0232401.
PR 14-SEP-2000; 2000US-0233063.
PR 14-SEP-2000; 2000US-0233064.
PR 14-SEP-2000; 2000US-0233065.
PR 21-SEP-2000; 2000US-0234223.
PR 21-SEP-2000; 2000US-0234274.
PR 25-SEP-2000; 2000US-0234997.
PR 25-SEP-2000; 2000US-0234998.
PR 26-SEP-2000; 2000US-0235484.
PR 27-SEP-2000; 2000US-0235834.
PR 27-SEP-2000; 2000US-0235836.
PR 29-SEP-2000; 2000US-0236327.
PR 29-SEP-2000; 2000US-0236367.
PR 29-SEP-2000; 2000US-0236368.
PR 29-SEP-2000; 2000US-0236369.
```





CC AAY36754-Y37949 are encoded by open reading frames (ORFs) of the genome  
CC of Chlamydia trachomatis (see AA201425). The polypeptides can be used as  
CC vaccines against Chlamydia trachomatis. Antisense and ribozyme sequences  
CC can also be used to control growth of the microorganism. Chlamydia  
CC trachomatis is responsible for a large number of diseases, e.g. eye  
CC diseases such as conventional trachoma, nonendemic trachoma,  
CC paratrachoma, and inclusion conjunctivitis; genital diseases such as  
CC nongonococcal urethritis, epididymitis, cervicitis, salpingitis,  
CC perihepatitis, Bartholinitis; pneumopathy in breast feeding infants;  
CC and venereal lymphogranulomatosis. The polypeptides of the invention  
CC may be of use in treating these diseases.

XX Sequence 91 AA;

Query Match 27.6%; Score 42; DB 20; Length 91;  
Best Local Similarity 56.2%; Pred. No. 66;  
Matches 9; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 7 LMANNIISKKEIKW 22  
Db 25 LEKQIDRNKKEISW 40

RESULT 37

ABG49671  
ID ABG49671 standard; Peptide; 75 AA.

XX AC ABG49671;

XX DT 25-FEB-2003 (first entry)

XX DE Human liver peptide, SEQ ID No 28319.

XX KW Human; liver; cirrhosis; hyperlipoproteinaemia; hyperlipidaemia;  
XX KW hypercholesterolaemia; coronary heart disease.

XX OS Homo sapiens.

XX PN WO200157273-A2.

XX PD 09-AUG-2001.

XX PF 30-JAN-2001; 2001WO-US000664.

XX PR 04-FEB-2000; 2000US-0180312.

XX PR 26-MAY-2000; 2000US-0207456.

XX PR 30-JUN-2000; 2000US-0608408.

XX PR 03-AUG-2000; 2000US-0632366.

XX PR 21-SEP-2000; 2000US-0234687.

XX PR 27-SEP-2000; 2000US-0236359.

XX PR 04-OCT-2000; 2000GB-0024263.

XX PA (MOLE-) MOLECULAR DYNAMICS INC.

XX PI Penn SG, Hanzel DK, Chen W, Rank DR;

XX WPI; 2001-488898/53.

XX PT Human genome-derived single exon nucleic acid probes useful for  
XX PT analysing gene expression in human adult liver -

XX FS Claim 27; SEQ ID No 28319; 658pp; English.

XX The invention relates to a single exon nucleic acid probe (SEN) (I) for  
CC measuring human gene expression in a sample derived from human adult  
CC liver, comprising one of 13109 defined nucleotide sequences given in the  
CC specification (or complements/ fragments). The probe hybridises at high  
CC stringency to a nucleic acid molecule expressed in the human adult  
CC liver. (I) may be used for predicting, measuring and displaying gene  
CC expression in samples derived from human adult liver. The genes  
CC identified may be involved in genetic liver diseases such as cirrhosis,  
CC hyperlipoproteinaemia, hyperlipidaemia and hypercholesterolaemia which  
CC is associated with coronary heart disease. ABG47348-ABG59930 represent

CC human liver single exon encoded peptides of the invention.  
CC Note: The sequence information for this patent does not appear in the  
CC printed specification but was obtained in electronic format directly  
CC from WIPO at [ftp.wipo.int/pub/published\\_pct\\_sequences](http://ftp.wipo.int/pub/published_pct_sequences).

XX Sequence 75 AA;

Query Match 27.3%; Score 41.5; DB 22; Length 75;  
Best Local Similarity 40.0%; Pred. No. 63;  
Matches 8; Conservative 4; Mismatches 3; Indels 5; Gaps 1;

QY 13 ISKEKEIKWIG-----LPT 27  
Db 31 ITKEGSLRWAGKFEIPT 50

RESULT 38

ABB29659  
ID ABB29659 standard; Peptide; 75 AA.

XX AC ABB29659;

XX DT 01-FEB-2002 (first entry)

XX DE Peptide #2310 encoded by breast cell single exon nucleic acid probe.

XX KW Human; microarray; single exon probe; gene expression; breast;  
XX KW disease; cancer.

XX OS Homo sapiens.

XX PN WO200157271-A2.

XX PD 09-AUG-2001.

XX PF 30-JAN-2001; 2001WO-US000662.

XX PR 04-FEB-2000; 2000US-0180312.

XX PR 26-MAY-2000; 2000US-0207456.

XX PR 30-JUN-2000; 2000US-0608408.

XX PR 03-AUG-2000; 2000US-0632366.

XX PR 21-SEP-2000; 2000US-0234687.

XX PR 27-SEP-2000; 2000US-0236359.

XX PR 04-OCT-2000; 2000GB-0024263.

XX PA (MOLE-) MOLECULAR DYNAMICS INC.

XX PI Penn SG, Hanzel DK, Chen W, Rank DR;

XX WPI; 2001-496933/54.

XX PT New spatially-addressable set of single exon nucleic acid probes,  
XX PT useful for measuring gene expression in sample derived from human  
XX PT breast, comprises number of single exon nucleic acid probes -

XX PS Claim 27; SEQ ID No 12627; 327pp + sequence listing; English.

XX The invention relates to a spatially-addressable set of single exon  
CC nucleic acid probes for measuring gene expression in a sample derived  
CC from human breast and BT 474 cells. The method involves contacting  
CC the probes with a collection of detectably labelled nucleic acids  
CC derived from mRNA of human breast, and then measuring the label  
CC bound to each probe of the microarray. The probes are useful for  
CC verifying the expression of regions of genomic DNA predicted to  
CC encode proteins. They are useful for gene discovery, and for  
CC determining predisposition and/or prognosing breast disease. Gene  
CC expression analysis is useful for assessing the toxicity of chemical  
CC agents on cells. The microarray of this invention presents a far greater  
CC diversity of probes for measuring gene expression, with far less bias  
CC than expressed sequence tag microarrays. The method is suitable for  
CC rapid production of functional information from genomic sequence. The  
CC present sequence is a peptide encoded by a single exon nucleic acid  
CC probe of the invention.

Tue Feb 17 11:55:54 2004

CC Note: The sequence data for this patent did not form part of the  
CC printed specification, but was obtained in electronic format directly  
CC from WIPO at ftp.wipo.int/pub/published\_pct\_sequences.

XX SQ Sequence 75 AA;  
XX Query Match 27.3%; Score 41.5; DB 22; Length 75;  
XX Best Local Similarity 40.0%; Pred. No. 63;  
XX Matches 8; Conservative 3; Indels 5; Gaps 1;

Qy 13 ISKEKEIKWIG-----LPT 27  
Db 31 ITKEKSLRWAGKKFEIPT 50

RESULT 39  
ABB34837  
ID ABB34837 standard; Peptide; 75 AA.

XX AC ABB34837;

XX DT 04-FEB-2002 (first entry)

XX DE Peptide #2343 encoded by human foetal liver single exon probe.

XX KW Human; foetal liver; gene expression; single exon nucleic acid probe.

XX OS Homo sapiens.

XX WO200157277-A2.

XX PN 09-AUG-2001.

XX PF 30-JAN-2001; 2001WO-US00669.

XX PR 04-FEB-2000; 2000US-0180312.

XX PR 26-MAY-2000; 2000US-0207456.

XX PR 30-JUN-2000; 2000US-0608408.

XX PR 03-AUG-2000; 2000US-0632366.

XX PR 21-SEP-2000; 2000US-0234687.

XX PR 27-SEP-2000; 2000US-0236359.

XX PR 04-OCT-2000; 2000GB-0024263.

XX PA (MOLE-) MOLECULAR DYNAMICS INC.

XX PI Penn SG, Hanzel DK, Chen W, Rank DR;

XX WPI; 2001-483447/52.

XX Human genome-derived single exon nucleic acid probes useful for

XX analyzing gene expression in human fetal liver -

XX Claim 27; SEQ ID NO 27472; 639pp + sequence listing; English.

XX The invention relates to a single exon nucleic acid probe for  
XX measuring human gene expression in a sample derived from human foetal  
XX liver. The single exon nucleic acid probes may be used for predicting,  
XX measuring and displaying gene expression in samples derived from human  
XX fetal liver. The present sequence is a peptide encoded by a single exon  
XX nucleic acid probe of the invention.

XX Note: The sequence data for this patent did not form part of the  
XX printed specification, but was obtained in electronic format directly  
XX from WIPO at ftp.wipo.int/pub/published\_pct\_sequences.

XX SQ Sequence 75 AA;

XX Query Match 27.3%; Score 41.5; DB 22; Length 75;

XX Best Local Similarity 40.0%; Pred. No. 63;

XX Matches 8; Conservative 3; Indels 5; Gaps 1;

Qy 13 ISKEKEIKWIG-----LPT 27  
Db 31 ITKEKSLRWAGKKFEIPT 50

RESULT 40

ABB20257

ID ABB20257 standard; Protein; 75 AA.

XX AC ABB20257;

XX DT 23-JAN-2002 (first entry)

XX DE Protein #2256 encoded by probe for measuring heart cell gene expression.

XX KW Human; gene expression; heart; microarray; vascular system;

XX KW cardiovascular disease; hypertension; cardiac arrhythmia;

XX KW congenital heart disease.

XX OS Homo sapiens.

XX WO200157274-A2.

XX PN 09-AUG-2001.

XX PF 30-JAN-2001; 2001WO-US00666.

XX PR 04-FEB-2000; 2000US-0180312.

XX PR 26-MAY-2000; 2000US-0207456.

XX PR 30-JUN-2000; 2000US-0608408.

XX PR 03-AUG-2000; 2000US-0632366.

XX PR 21-SEP-2000; 2000US-0234687.

XX PR 27-SEP-2000; 2000US-0236359.

XX PR 04-OCT-2000; 2000GB-0024263.

XX PA (MOLE-) MOLECULAR DYNAMICS INC.

XX PI Penn SG, Hanzel DK, Chen W, Rank DR;

XX WPI; 2001-488999/53.

XX Single exon nucleic acid probes for analyzing gene expression in human

XX hearts -

XX Claim 15; SEQ ID No 22027; 530pp; English.

XX The present invention relates to single exon nucleic acid probes for  
XX measuring human gene expression in a sample derived from human heart (see  
XX ABA21535-ABA41305). The present sequence is a protein encoded by one such  
XX probe. The probes may be used for predicting, measuring and displaying  
XX gene expression in samples derived from the human heart via microarrays.  
XX By measuring gene expression, the probes are useful for predicting,  
XX diagnosing, grading, staging, monitoring and prognosing diseases of the  
XX human heart and vascular system e.g. cardiovascular disease,  
XX hypertension, cardiac arrhythmias and congenital heart disease.

XX Note: The sequence data for this patent did not form part of the printed  
XX specification, but was obtained in electronic format directly from WIPO  
XX at ftp.wipo.int/pub/published\_pct\_sequences.

XX SQ Sequence 75 AA;

XX Query Match 27.3%; Score 41.5; DB 22; Length 75;

XX Best Local Similarity 40.0%; Pred. No. 63;

XX Matches 8; Conservative 3; Indels 5; Gaps 1;

Qy 13 ISKEKEIKWIG-----LPT 27  
Db 31 ITKEKSLRWAGKKFEIPT 50

RESULT 41

AAM55645

ID AAM55645 standard; Protein; 75 AA.

XX AC AAM55645;

XX AC AAM55645;



Tue Feb 17 11:55:54 2004

us-09-900-147-6.rag

CC from human placenta. The probes are useful for antenatal diagnosis of  
CC human genetic disorders.

XX Sequence 75 AA;

Query Match 27.3%; Score 41.5; DB 22; Length 75;

Best Local Similarity 40.0%; Pred. No. 63;

Matches 8; Conservative 3; Indels 5; Gaps 1;

QY 13 ISKEKEIKWIG-----LPT 27

DB 31 ITKEKSLRWAGKPFEBIPT 50

RESULT 45

AA03581

ID AA03581 standard; Protein; 75 AA.

XX AC AA03581;

XX DT 09-OCT-2001 (first entry)

XX DE Peptide #2263 encoded by probe for measuring breast gene expression.

XX KW Probe; human; breast disease; breast cancer; development disorder;

XX KW inflammatory disease; proliferative breast disease; non-carcinoma tumour.

XX OS Homo sapiens.

XX PN WO200157270-A2.

XX PD 09-AUG-2001.

XX PF 29-JAN-2001; 2001WO-US00661.

XX PR 04-FEB-2000; 2000US-0180312.

XX PR 26-MAY-2000; 2000US-0207456.

XX PR 30-JUN-2000; 2000US-0608408.

XX PR 03-AUG-2000; 2000US-0632366.

XX PR 21-SEP-2000; 2000US-0234687.

XX PR 27-SEP-2000; 2000US-0236359.

XX PR 04-OCT-2000; 2000GB-0024263.

XX PA (MOLE-) MOLECULAR DYNAMICS INC.

XX PI Penn SG, Hanzel DK, Chen W, Rank DR;

XX XX WPI; 2001-476286/51.

XX DR Novel single exon nucleic acid probe used to measuring gene expression  
in a human breast -

XX PS Claim 27; SEQ ID No 12321; 322pp; English.

XX CC The present invention relates to novel single exon nucleic acid probes  
(see AA100010-AA110067). The present sequence is a peptide encoded by one  
such probe. The probes are useful for measuring human gene expression in  
a human breast sample, where the probe hybridises at high stringency to a  
nucleic acid expressed in the human breast. The probes are useful for  
predicting, diagnosing, grading, staging, monitoring and prognosing  
diseases of the human breast, particularly those diseases with polygenic  
aetiology. The diseases include: breast cancer, disorders of development,  
inflammatory diseases of the breast, fibrocystic changes, proliferative  
breast disease and non-carcinoma tumours.

XX CC Note: The sequence data for this patent did not form part of the printed  
specification, but was obtained in electronic format directly from WIPO  
at ftp.wipo.int/pub/published\_pct\_sequences.

XX SQ Sequence 75 AA;

Query Match 27.3%; Score 41.5; DB 22; Length 75;

Best Local Similarity 40.0%; Pred. No. 63;

Matches 8; Conservative 4; Mismatches 3; Indels 5; Gaps 1;

XX Human genome-derived single exon nucleic acid probes useful for

PT analyzing gene expression in human cervical epithelial cells -

XX Claim 27; SEQ ID No 20669; 487pp; English.

XX CC The present invention relates to human single exon nucleic acid probes  
(SENPs: see AA110068-AA128459). The present sequence is a peptide encoded  
by one such probe. The SENPs are derived from human Hela cells. The SENPs  
can be used to produce a single exon microarray, which can be used for  
measuring human gene expression in a sample derived from human cervical  
epithelial cells. By measuring gene expression, the probes are therefore  
useful in grading and/or staging of diseases of the cervix, notably  
cervical cancer.

XX CC Note: The sequence data for this patent did not form part of the printed  
specification, but was obtained in electronic format directly from WIPO  
at ftp.wipo.int/pub/published\_pct\_sequences.

XX SQ Sequence 75 AA;

Query Match 27.3%; Score 41.5; DB 22; Length 75;

Best Local Similarity 40.0%; Pred. No. 63;

Matches 8; Conservative 4; Mismatches 3; Indels 5; Gaps 1;

QY 13 ISKEKEIKWIG-----LPT 27

DB 31 ITKEKSLRWAGKPFEBIPT 50

RESULT 44

AA028353

ID AA028353 standard; Protein; 75 AA.

XX AC AA028353;

XX DT 17-OCT-2001 (first entry)

XX DE Peptide #2390 encoded by probe for measuring placental gene expression.

XX KW Probe; microarray; human; placenta; antenatal diagnosis;

XX KW Genetic disorder.

XX OS Homo sapiens.

XX PN WO200157272-A2.

XX PD 09-AUG-2001.

XX PF 30-JAN-2001; 2001WO-US00663.

XX PR 04-FEB-2000; 2000US-0180312.

XX PR 26-MAY-2000; 2000US-0207456.

XX PR 30-JUN-2000; 2000US-0608408.

XX PR 03-AUG-2000; 2000US-0632366.

XX PR 21-SEP-2000; 2000US-0234687.

XX PR 27-SEP-2000; 2000US-0236359.

XX PR 04-OCT-2000; 2000GB-0024263.

XX PA (MOLE-) MOLECULAR DYNAMICS INC.

XX PI Penn SG, Hanzel DK, Chen W, Rank DR;

XX XX WPI; 2001-488997/53.

XX DR Human genome-derived single exon nucleic acid probes useful for  
analyzing gene expression in human placenta -

XX PS Claim 27; SEQ ID No 28622; 654pp; English.

XX CC The present invention relates to single exon nucleic acid probes (SENPs:  
see AA131315-AA157546). The present sequence is a peptide encoded by one  
such probe. The probes are useful for producing a microarray for  
predicting, measuring and displaying gene expression in samples derived

Search completed: February 17, 2004, 10:53:43  
Job time : 58.0297 secs

**THIS PAGE BLANK (USPTO)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: February 17, 2004, 10:50:12 ; Search time 13.3069 Seconds  
(without alignments)  
83.497 Million cell updates/sec

Title: US-09-900-147-7

Perfect score: 31

Sequence: 1 ALNVLMA 7

Scoring table:

Gapop 10.0 , Gapext 0.5

Searched: 1107863 seqs, 158726573 residues

Total number of hits satisfying chosen parameters: 717921

Minimum DB seq length: 0

Maximum DB seq length: 100

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : A Geneseq 19Jun03.\*

1: /SIDSL1/gcgdata/geneseq/geneseq-emb1/AA1980.DAT.\*  
2: /SIDSL1/gcgdata/geneseq/geneseq-emb1/AA1981.DAT.\*  
3: /SIDSL1/gcgdata/geneseq/geneseq-emb1/AA1982.DAT.\*  
4: /SIDSL1/gcgdata/geneseq/geneseq-emb1/AA1983.DAT.\*  
5: /SIDSL1/gcgdata/geneseq/geneseq-emb1/AA1984.DAT.\*  
6: /SIDSL1/gcgdata/geneseq/geneseq-emb1/AA1985.DAT.\*  
7: /SIDSL1/gcgdata/geneseq/geneseq-emb1/AA1986.DAT.\*  
8: /SIDSL1/gcgdata/geneseq/geneseq-emb1/AA1987.DAT.\*  
9: /SIDSL1/gcgdata/geneseq/geneseq-emb1/AA1988.DAT.\*  
10: /SIDSL1/gcgdata/geneseq/geneseq-emb1/AA1989.DAT.\*  
11: /SIDSL1/gcgdata/geneseq/geneseq-emb1/AA1990.DAT.\*  
12: /SIDSL1/gcgdata/geneseq/geneseq-emb1/AA1991.DAT.\*  
13: /SIDSL1/gcgdata/geneseq/geneseq-emb1/AA1992.DAT.\*  
14: /SIDSL1/gcgdata/geneseq/geneseq-emb1/AA1993.DAT.\*  
15: /SIDSL1/gcgdata/geneseq/geneseq-emb1/AA1994.DAT.\*  
16: /SIDSL1/gcgdata/geneseq/geneseq-emb1/AA1995.DAT.\*  
17: /SIDSL1/gcgdata/geneseq/geneseq-emb1/AA1996.DAT.\*  
18: /SIDSL1/gcgdata/geneseq/geneseq-emb1/AA1997.DAT.\*  
19: /SIDSL1/gcgdata/geneseq/geneseq-emb1/AA1998.DAT.\*  
20: /SIDSL1/gcgdata/geneseq/geneseq-emb1/AA1999.DAT.\*  
21: /SIDSL1/gcgdata/geneseq/geneseq-emb1/AA2000.DAT.\*  
22: /SIDSL1/gcgdata/geneseq/geneseq-emb1/AA2001.DAT.\*  
23: /SIDSL1/gcgdata/geneseq/geneseq-emb1/AA2002.DAT.\*  
24: /SIDSL1/gcgdata/geneseq/geneseq-emb1/AA2003.DAT.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	31	100.0	7	AAW30503	DP-1 transcription
2	31	100.0	14	AAW30511	DP-1 transcription
3	31	100.0	15	AAW57052	E2F activity inhib
4	31	100.0	16	AAW30506	DP-1 transcription
5	31	100.0	19	AAW30516	DP-1 transcription
6	31	100.0	19	AAW30504	DP-1 transcription
7	31	100.0	28	AAW57051	E2F activity inhib
8	31	100.0	28	AAW57055	E2F activity inhib
9	31	100.0	30	AAW30507	DP-1 transcription

10	31	100.0	37	19	AAW30501	DP-1 transcription
11	31	100.0	83	20	AAW32163	Soybean DP-1 prote
12	27	87.1	8	23	AAU72578	Plant dimerisation
13	27	87.1	19	19	AAW30515	DP-1 transcription
14	27	87.1	56	22	AAW60781	Scorpion leuropep
15	26	83.9	35	23	AAU72601	DEF domain consens
16	26	83.9	41	21	AAW08901	Human secreted pro
17	26	83.9	90	21	AAW53284	Human colon cancer
18	25	80.6	59	22	AAW05442	Human polyepitide
19	25	80.6	73	22	AAW24066	Human EST encoded
20	25	80.6	87	19	AAW75232	Human secreted pro
21	25	80.6	87	23	AAE26996	Human gene 20 enco
22	25	80.6	87	23	AAE27134	Human gene 20 enco
23	25	80.6	87	24	ABU65007	Human secreted pro
24	24	77.4	25	19	AAW69785	Human iduronate-2-
25	24	77.4	25	19	AAW47300	Human IDS peptide
26	24	77.4	25	20	AAW23983	Amino terminal seq
27	24	77.4	25	22	AAW51178	Human iduronate 2
28	24	77.4	27	22	ABG49596	Human liver peptid
29	24	77.4	27	22	ABE29588	Peptide #2239 enco
30	24	77.4	27	22	ABB34768	Peptide #2274 enco
31	24	77.4	27	22	ABB20183	Protein #2182 enco
32	24	77.4	27	22	AAW55570	Human brain expres
33	24	77.4	27	22	AAW67954	Human bone marrow
34	24	77.4	27	22	AAW15772	Peptide #2206 enco
35	24	77.4	27	22	AAW28281	Peptide #2118 enco
36	24	77.4	27	22	AAW03506	Peptide #2188 enco
37	24	77.4	27	23	ABG37489	Human peptide enco
38	24	77.4	32	22	AAW77060	Human colon cancer
39	24	77.4	61	23	ABP10982	Human ORFX protein
40	24	77.4	64	22	AAW91228	C glutamicum prote
41	24	77.4	65	22	AAU17920	Novel human respir
42	24	77.4	71	21	AAW56878	Human prostate can
43	24	77.4	75	22	AAU27363	Novel bone marrow
44	23	74.2	7	22	AAW09109	Bacteriophage Type
45	23	74.2	9	19	AAW30502	DP-1 transcription

#### ALIGNMENTS

##### RESULT 1

AAW30503  
ID AAW30503 standard; Peptide; 7 AA.

XX AC AAW30503;

XX DT 26-OCT-1998 (first entry)

XX DE DP-1 transcription factor antagonist peptide motif.

XX DE DP-1; transcription factor; antagonist; E2F protein; apoptosis;

KW cell proliferation; cardiovascular cell; restenosis; tumour;

KW surgical stent; therapy.

XX OS Synthetic.

XX OS Homo sapiens.

XX OS Homo sapiens.

PN WO9828334-A1.

PD 02-JUL-1998.

XX 22-DEC-1997; 97WO-GB03506.

XX 20-DEC-1996; 96GB-0026589.

XX (PROL-) PROLIFIX LTD.

XX Bandara LR, La Thangue NB;

XX WPI; 1998-377596/32.

XX Polypeptide fragments of the DP-1 transcription factor - used for

Tue Feb 17 11:55:55 2004

PT inducing apoptosis, specifically in tumour and cardiovascular cells,  
 PT e.g. for preventing restenosis,  
 XX  
 PS Claim 3; Page 44; 55pp; English.  
 XX This peptide comprises amino acid residues 172-178 in the DEF box  
 CC region (see AAW30501) of transcription factor Dp1. Claimed peptides  
 CC (II) (see AAW30502-07) containing this and/or another motif (see  
 CC AAW30502) of the DEF box, are antagonists of the heterodimerisation  
 CC of a Dp protein with an E2F protein. Also claimed are variants  
 CC of these peptides, especially containing substitutions of residues  
 CC corresponding to residues 167, 169, 171 and 175 of Dp-1, fusion  
 CC proteins (III) comprising (I) or (II) and a membrane translocation  
 CC sequence (see AAW30508), expression vectors encoding (I)-(III) and  
 CC host cells. (I)-(III) are used therapeutically to induce apoptosis,  
 CC specifically in tumour or cardiovascular cells, either in vivo or in  
 CC vitro, e.g. for purging bone marrow. Surgical stents comprising  
 CC (I)-(III) are used to treat or prevent restenosis in patients who  
 CC have undergone angioplasty. (I)-(III) function by inactivating  
 CC the DNA-binding activity of Dp/E2F heterodimers. They are also  
 CC used as research reagents, as positive controls in assays for  
 CC identifying antagonists of Dp-1/E2F dimerisation and as immunoassay  
 CC agents. Also described is the use of sequences antisense to  
 CC nucleic acids encoding (I)-(III) to control Dp levels in cells,  
 CC particularly by gene therapy. When formulated with cytotoxic  
 CC or cytostatic agents, (I)-(III) enhance cell killing.  
 XX

XX Sequence 7 AA;  
 SQ  
 Query Match 100.0%; Score 31; DB 19; Length 7;  
 Best Local Similarity 100.0%; Pred. No. 9.3e+05;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ALNVLMA 7  
 DB 1 ALNVLMA 7

RESULT 2  
 AAW30511  
 ID AAW30511 standard; Peptide; 14 AA.  
 AC  
 XX AAW30511;  
 XX

26-OCT-1998 (first entry)

DE DP-1 transcription factor peptide H6.

XX DP-1; transcription factor; antagonist; E2F protein; apoptosis;  
 KW cell proliferation; cardiovascular cell; restenosis; tumour;  
 KW surgical stent; therapy.  
 KW

OS Synthetic.  
 OS Homo sapiens.

FN WO9828334-A1.

XX 02-JUL-1998.

XX 22-DEC-1997; 97WO-GB03506.

XX 20-DEC-1996; 96GB-0026589.

PA (PROL-) PROLIFIX LTD.

XX Bandara LR, La Thangue NB;

XX WPI; 1998-377596/32.

XX Polypeptide fragments of the Dp-1 transcription factor - used for  
 PT inducing apoptosis, specifically in tumour and cardiovascular cells,  
 PT e.g. for preventing restenosis  
 XX

PS Example C; Page 41; 55pp; English.

XX Peptide H6 comprises amino acid residues 167-180 in the DEF box  
 CC region (see AAW30501) of transcription factor Dp1. Unlike claimed  
 CC peptides (see AAW30504-07) that contain one or both of 2 motifs (see  
 CC AAW30502-03) of the Dp1 DEF box, peptide H6 is not capable of  
 CC antagonising the heterodimerisation of a Dp protein with an E2F  
 CC protein. The claimed peptides, their variants and fusion proteins  
 CC can be used to induce apoptosis, specifically in tumour and  
 CC cardiovascular cells, e.g. to prevent restenosis.

XX Sequence 14 AA;

Query Match 100.0%; Score 31; DB 19; Length 14;

Best Local Similarity 100.0%; Pred. No. 0.69;

Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ALNVLMA 7  
 DB 6 ALNVLMA 12

RESULT 3

AAW57052

ID AAW57052 standard; peptide; 15 AA.

XX AAW57052;

XX 28-AUG-1998 (first entry)

XX E2F activity inhibiting compound Ib-2.

DE E2F activity; inhibitor; treatment; tumour; arteriosclerosis.

OS Synthetic.

OS Homo sapiens.

XX Key Location/Qualifiers

FT Modified-site 1 /note= "N-terminal acetyl"

FT Modified-site 15 /note= "C-terminal amide"

XX WO9814474-A1.

XX 09-APR-1998.

XX 26-SEP-1997; 97WO-JP03442.

XX 30-SEP-1996; 96JP-0259432.

XX (KYOW ) KYOWA HAKKO KOGYO KK.

XX Mizukami T, Shibata K, Yamasaki M, Yoshida T;

XX WPI; 1998-240020/21.

XX E2F activity inhibitors - for treatment and prevention of tumours  
 and arteriosclerosis  
 XX

PS Example 4; Page 28; 52pp; Japanese.

XX This represents a compound that can inhibit E2F activity. The compound  
 CC is of the formula R1 - A - R2 where R1 is an optionally substituted  
 CC alkanoyl, allyl, hetero-arylcabonyl, alkoxycarbonyl, aryloxy carbonyl,  
 CC hetero-aryloxy carbonyl, or H, R2 is OH, or optionally substituted alkoxy  
 CC or amino, and A is an E2F family dimer forming region or DNA binding  
 CC region, of at least 12 consecutive amino acids. Compounds of this formula  
 CC can be used to inhibit E2F activity, and are useful in the treatment and  
 CC prevention of tumours and arteriosclerosis.

XX Sequence 15 AA;

XX SQ



Query Match 100.0%; Score 31; DB 19; Length 15;  
 Best Local Similarity 100.0%; Pred. No. 0.75;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ALNVLMA 7  
 Db 8 ALNVLMA 14

RESULT 4  
 AAW30506  
 ID AAW30506 standard; Peptide; 16 AA.

XX AC AAW30506;

XX DT 26-OCT-1998 (first entry)

XX DE DP-1 transcription factor antagonist peptide H5.

XX KW DP-1; transcription factor; antagonist; E2F protein; apoptosis;  
 cell proliferation; cardiovascular cell; restenosis; tumour;  
 KW surgical stent; therapy.

XX OS Synthetic.  
 OS Homo sapiens.

XX FH Key Location/Qualifiers  
 FT Peptide 5..11  
 FT /note= "Claim 3"  
 FT Peptide 7..16  
 FT /note= "Claim 3"

XX PN WO9828334-A1.

XX PD 02-JUL-1998.

XX PF 22-DEC-1997; 97WO-GB03506.

XX PR 20-DEC-1996; 96GB-0026589.

XX PA (PROL-) PROLIFIX LTD.

XX PI Bandara LR, La Thangue NB;

XX DR WPI; 1998-377596/32.

XX PT Polypeptide fragments of the DP-1 transcription factor - used for  
 inducing apoptosis, specifically in tumour and cardiovascular cells,  
 PT e.g. for preventing restenosis

XX PS Claim 4; Page 44; 55pp; English.

XX CC Peptide H5 comprises amino acid residues 168-183 in the DEF box  
 CC (I) (see AAW30501) of transcription factor Dp1. Claimed peptides  
 CC (II) (see AAW30504-07) containing one or both of 2 motifs (see  
 CC AAW30502-03) of the DEF box are capable of antagonising the  
 CC heterodimerisation of a DP protein with an E2F protein. Also  
 CC claimed are variants of these peptides, especially containing  
 CC substitutions of residues corresponding to residues 167, 169, 171  
 CC and 175 of DP-1, fusion proteins (III) comprising (I) or (II) and a  
 CC membrane translocation sequence (see AAW30508), expression vectors  
 CC encoding (I)-(III) and host cells. (I)-(III) are used  
 CC therapeutically to induce apoptosis, specifically in tumour or  
 CC cardiovascular cells, either in vivo or in vitro, e.g. for purging  
 CC bone marrow. Surgical stents comprising (I)-(III) are used to  
 CC treat or prevent restenosis in patients who have undergone  
 CC angioplasty. (I)-(III) function by inactivating the DNA-binding  
 CC activity of DP/E2F heterodimers. They are also used as research  
 CC reagents, as positive controls in assays for identifying  
 CC antagonists of DP-1/E2F dimerisation and as immunoassay agents.  
 CC Also described is the use of sequences antisense to nucleic acids  
 CC encoding (I)-(III) to control DP levels in cells, particularly by  
 CC gene therapy. When formulated with cytotoxic or cytostatic agents,

CC (I)-(III) enhance cell killing.

XX SQ Sequence 16 AA;

Query Match 100.0%; Score 31; DB 19; Length 16;  
 Best Local Similarity 100.0%; Pred. No. 0.81;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ALNVLMA 7  
 Db 5 ALNVLMA 11

RESULT 5  
 AAW30516

ID AAW30516 standard; Peptide; 19 AA.

XX AC AAW30516;

XX DT 26-OCT-1998 (first entry)

XX DE DP-1 transcription factor antagonist peptide H2mt2.

XX KW DP-1; transcription factor; E2F protein; apoptosis;  
 cell proliferation; cardiovascular cell; restenosis; tumour;  
 KW surgical stent; therapy.

XX OS Synthetic.  
 OS Homo sapiens.

XX FH Key Location/Qualifiers  
 FT Misc-difference 2 /note= "R167A mutation"  
 FT Misc-difference 6 /note= "D171A mutation"

XX PN WO9828334-A1.

XX PD 02-JUL-1998.

XX PF 22-DEC-1997; 97WO-GB03506.

XX PR 20-DEC-1996; 96GB-0026589.

XX PA (PROL-) PROLIFIX LTD.

XX PI Bandara LR, La Thangue NB;

XX DR WPI; 1998-377596/32.

XX PT Polypeptide fragments of the DP-1 transcription factor - used for  
 inducing apoptosis, specifically in tumour and cardiovascular cells,  
 PT e.g. for preventing restenosis

XX PS Example D; Page 26; 55pp; English.

XX CC Peptide H2mt2 is based on peptide H2 (see AAW30504) from the DEF box  
 CC (see AAW30501) of transcription factor Dp1. In H2mt2, amino acid  
 CC residues of H2 that correspond to Dp1 residues Arg167 and Asp171  
 CC are substituted by Ala residues. H2 is an antagonist of the  
 CC heterodimerisation of Dp1 with E2F. H2mt2 retains some, but not  
 CC all, of this antagonistic activity. H2 and other claimed peptides  
 CC (see AAW30504-07) from the DEF box region of Dp1 can be used to  
 CC induce apoptosis, specifically in tumour and cardiovascular cells,  
 CC e.g. for the prevention of restenosis.

XX SQ Sequence 19 AA;

Query Match 100.0%; Score 31; DB 19; Length 19;  
 Best Local Similarity 100.0%; Pred. No. 0.99;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ALNVLMA 7

Tue Feb 17 11:55:55 2004

Db 1 ALNVLMMA 7  
 7 ALNVLMMA 13

## RESULT 6

AAW30504  
 ID AAW30504 standard; Peptide; 19 AA.

XX AC AAW30504;

XX DT 26-OCT-1998 (first entry)

XX DE DP-1 transcription factor antagonist peptide H2.

XX DE DP-1; transcription factor; antagonist; E2F protein; apoptosis;  
 KW cell proliferation; cardiovascular cell, restenosis; tumour;  
 KW surgical stent; therapy.

XX OS Synthetic.

XX OS Homo sapiens.

XX FH Key Location/Qualifiers  
 FT Peptide 9..18  
 FT /note= "Claim 3"

XX FT WO9828334-A1.

XX PN 02-JUL-1998.

XX PD 22-DEC-1997; 97WO-GB03506.

XX PF 20-DEC-1996; 96GB-0026589.

XX PR (PROL-) PROLIFIX LTD.

XX PI Bandara LR, La Thangue NB;

XX PI WPI; 1998-377596/32.

XX PT Polypeptide fragments of the DP-1 transcription factor - used for  
 PT inducing apoptosis, specifically in tumour and cardiovascular cells,  
 PT e.g. for preventing restenosis

XX PS Claim 4; Page 44; 55pp; English.

XX CC Peptide H2 comprises amino acid residues 166-184 in the DEF box  
 CC (I) (see AAW30501) of transcription factor DP1. Claimed peptides  
 CC (II) (see AAW30504-07) containing one or both of 2 motifs (see  
 CC AAW30502-03) of the DEF box are capable of antagonising the  
 CC heterodimerisation of a DP protein with an E2F protein. Also  
 CC claimed are variants of these peptides, especially containing  
 CC substitutions of residues corresponding to residues 167, 169, 171  
 CC and 175 of DP-1, fusion proteins (III) comprising (I) or (II) and a  
 CC membrane translocation sequence (see AAW30508), expression vectors  
 CC encoding (I)-(III) and host cells. (I)-(III) are used  
 CC therapeutically to induce apoptosis, specifically in tumour or  
 CC cardiovascular cells, either in vivo or in vitro, e.g. for purging  
 CC bone marrow. Surgical stents comprising (I)-(III) are used to  
 CC treat or prevent restenosis in patients who have undergone  
 CC angioplasty. (I)-(III) function by inactivating the DNA-binding  
 CC activity of DP/E2F heterodimers. They are also used as research  
 CC reagents, as positive controls in assays for identifying  
 CC antagonists of DP-1/E2F dimerisation and as immunoassay agents.  
 CC Also described is the use of sequences antisense to nucleic acids  
 CC encoding (I)-(III) to control DP levels in cells, particularly by  
 CC gene therapy. When formulated with cytotoxic or cytostatic agents,  
 CC (I)-(III) enhance cell killing.

XX SQ Sequence 19 AA;

Query Match 100.0%; Score 31; DB 19; Length 19;

Best Local Similarity 100.0%; Pred. No. 0.99; 0; Indels 0; Gaps 0;  
 Matches 7; Conservative 0; Mismatches 0;

QY 1 ALNVLMMA 7  
 Db 7 ALNVLMMA 13

## RESULT 7

AAW57051  
 ID AAW57051 standard; peptide; 28 AA.

XX AC AAW57051;

XX DT 28-AUG-1998 (first entry)

XX DE E2F activity inhibiting compound Ib-1.

XX DE E2F activity; inhibitor; treatment; tumour; arteriosclerosis.

XX OS Synthetic.

XX OS Homo sapiens.

XX FH Key Location/Qualifiers  
 FT Modified-site 1  
 FT /note= "N-terminal acetyl"  
 FT Modified-site 28  
 FT /note= "C-terminal amide"

XX PN WO9814474-A1.

XX PD 09-APR-1998.

XX PF 26-SEP-1997; 97WO-JP03442.

XX PR 30-SEP-1996; 96JP-0259432.

XX PA (KYOW ) KYOWA HAKKO KOGYO KK.

XX PI Mizukami T, Shibata K, Yamasaki M, Yoshida T;

XX PI WPI; 1998-240020/21.

XX PT E2F activity inhibitors - for treatment and prevention of tumours  
 PT and arteriosclerosis

XX PS Example 3; Page 27; 52pp; Japanese.

XX CC This represents a compound that can inhibit E2F activity. The compound  
 CC is of the formula R1 - A - R2 where R1 is an optionally substituted  
 CC alkanoyl, allyl, hetero-arylcabonyl, alkoxycarbonyl, aryloxyacetyl,  
 CC hetero-aryloxyacetyl, or H, R2 is OH, or optionally substituted alkoxy  
 CC or amino, and A is an E2F family dimer forming region or DNA binding  
 CC region, of at least 12 consecutive amino acids. Compounds of this formula  
 CC can be used to inhibit E2F activity, and are useful in the treatment and  
 CC prevention of tumours and arteriosclerosis.

XX SQ Sequence 28 AA;

Query Match 100.0%; Score 31; DB 19; Length 28;

Best Local Similarity 100.0%; Pred. No. 1.6; 0; Indels 0; Gaps 0;  
 Matches 7; Conservative 0; Mismatches 0;

QY 1 ALNVLMMA 7  
 Db 17 ALNVLMMA 23

## RESULT 8

AAW57055  
 ID AAW57055 standard; peptide; 28 AA.

XX AC AAW57055;

XX DT 28-AUG-1998 (first entry)

XX DE E2F activity inhibiting compound Ib-3.  
 XX KW E2F activity; inhibitor; treatment; tumour; arteriosclerosis.  
 XX OS Synthetic.  
 XX OS Homo sapiens.  
 XX FH Key Location/Qualifiers  
 FT Modified-site 1  
 FT /note= "N-terminal lauroyl"  
 FT 28  
 FT Modified-site  
 FT /note= "C-terminal amide"  
 XX WO9814474-A1.  
 XX PN  
 XX PD 09-APR-1998.  
 XX PF 26-SEP-1997; 97WO-JP03442.  
 XX PR 30-SEP-1996; 96JP-0259432.  
 XX (KYOW ) KYOWA HAKKO KOGYO KK.  
 XX PA Mizukami T, Shibata K, Yamasaki M, Yoshida T;  
 XX WIPI; 1998-240020/21.  
 XX DR  
 XX PT E2F activity inhibitors - for treatment and prevention of tumours  
 XX PT and arteriosclerosis  
 XX PS Example 7; Page 33; 52pp; Japanese.  
 XX CC This represents a compound that can inhibit E2F activity. The compound  
 CC is of the formula R1 - A - R2 where R1 is an optionally substituted  
 CC alkanoyl, allyl, hetero-arylcarbonyl, alkoxycarbonyl, aryloxy carbonyl,  
 CC hetero-aryloxy carbonyl, or H, R2 is OH, or optionally substituted alkoxy  
 CC or amino, and A is an E2F family dimer forming region or DNA binding  
 CC region, of at least 12 consecutive amino acids. Compounds of this formula  
 CC can be used to inhibit E2F activity, and are useful in the treatment and  
 CC prevention of tumours and arteriosclerosis.  
 XX SQ Sequence 28 AA;  
 Query Match 100.0%; Score 31; DB 19; Length 28;  
 Best Local Similarity 100.0%; Pred. No. 1.6;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 ALNVLMA 7  
 Db 17 ALNVLMA 23  
 RESULT 9  
 AAW30507  
 ID AAW30507 standard; Peptide; 30 AA.  
 XX AC  
 XX AAW30507;  
 XX DT 26-OCT-1998 (first entry)  
 XX DE DP-1 transcription factor antagonist peptide H7.  
 XX DP-1; transcription factor; antagonist; E2F protein; apoptosis;  
 KW cell proliferation; cardiovascular cell; restenosis; tumour;  
 KW surgical stent; therapy.  
 XX OS Synthetic.  
 OS Homo sapiens.  
 XX Key Location/Qualifiers  
 FT Peptide 3..9  
 FT /note= "Claim 3"

FT Peptide 5..15  
 FT /note= "Claim 3"  
 XX WO9828334-A1.  
 XX PN 02-JUL-1998.  
 XX PD 22-DEC-1997; 97WO-GB03506.  
 XX PF 20-DEC-1996; 96GB-0026589.  
 XX PR (PROL-) PROLIFIX LTD.  
 XX PA Bandara LR, La Thangue NB;  
 XX PI WIPI; 1998-377596/32.  
 XX DR  
 XX PT Polypeptide fragments of the DP-1 transcription factor - used for  
 PT inducing apoptosis, specifically in tumour and cardiovascular cells,  
 PT e.g. for preventing restenosis.  
 XX PS Claim 4; Page 44; 55pp; English.  
 XX CC Peptide H7 comprises amino acid residues 170-199 in the DEF box  
 CC (I) (see AAW30501) of transcription factor DPl. Claimed peptides  
 CC (II) (see AAW30504-07) containing one or both of 2 motifs (see  
 CC AAW30502-03) of the DEF box are capable of antagonising the  
 CC heterodimerisation of a DP protein with an E2F protein. Also  
 CC claimed are variants of these peptides, especially containing  
 CC substitutions of residues corresponding to residues 167, 169, 171  
 CC and 175 of DP-1, fusion proteins (III) comprising (I) or (II) and a  
 CC membrane translocation sequence (see AAW30508), expression vectors  
 CC encoding (I)-(III) and host cells. (I)-(III) are used  
 CC therapeutically to induce apoptosis, specifically in tumour or  
 CC cardiovascular cells, either in vivo or in vitro, e.g. for purging  
 CC bone marrow. Surgical stents comprising (I)-(III) are used to  
 CC treat or prevent restenosis in patients who have undergone  
 CC angioplasty. (I)-(III) function by inactivating the DNA-binding  
 CC activity of DP/E2F heterodimers. They are also used as research  
 CC reagents, as positive controls in assays for identifying  
 CC antagonists of DP-1/E2F dimerisation and as immunoassay agents.  
 CC Also described is the use of sequences antisense to nucleic acids  
 CC encoding (I)-(III) to control DP levels in cells, particularly by  
 CC gene therapy. When formulated with cytotoxic or cytostatic agents,  
 CC (I)-(III) enhance cell killing.  
 XX SQ Sequence 30 AA;  
 Query Match 100.0%; Score 31; DB 19; Length 30;  
 Best Local Similarity 100.0%; Pred. No. 1.7;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 ALNVLMA 7  
 Db 3 ALNVLMA 9  
 RESULT 10  
 AAW30501  
 ID AAW30501 standard; Peptide; 37 AA.  
 XX AC  
 XX AAW30501;  
 XX DT 26-OCT-1998 (first entry)  
 XX DE DP-1 transcription factor peptide H (DEF box).  
 XX DP-1; transcription factor; antagonist; E2F protein; apoptosis;  
 KW cell proliferation; cardiovascular cell; restenosis; tumour;  
 KW surgical stent; therapy.  
 XX OS Synthetic.  
 OS Homo sapiens.

Tue Feb 17 11:55:55 2004

us-09-900-147-7.rag

XX PN WO9828334-A1.  
XX PD 02-JUL-1998.  
XX PF 22-DEC-1997; 97WO-GB03506.  
XX PR 20-DEC-1996; 96GB-0026589.  
XX PA (PROL-) PROLIFIX LTD.  
XX PI Bandara LR, La Thangue NB;  
XX DR WPI; 1998-377596/32.  
XX PT Polypeptide fragments of the DP-1 transcription factor - used for  
FT inducing apoptosis, specifically in tumour and cardiovascular cells,  
FT e.g. for preventing restenosis  
XX XX  
XX PS Claim 1; Page 44; 55pp; English.  
XX PN Peptide H (I) comprises residues 163-199, i.e. the DEF box region,  
CC of transcription factor Dp1. Claimed fragments (II) (see AAW30502-07)  
CC of (I) are capable of antagonising the heterodimerisation of a DP  
CC protein with an E2F protein. Also claimed are fusion proteins  
CC (III) comprising (I) or (II) and a membrane translocation sequence  
CC (see AAW30508), expression vectors encoding (I)-(III) and host cells.  
CC (I)-(III) are used therapeutically to induce apoptosis,  
CC specifically in tumour or cardiovascular cells, either in vivo or in  
CC vitro, e.g. for purging bone marrow. Surgical stents comprising  
CC (I)-(III) are used to treat or prevent restenosis in patients who  
CC have undergone angioplasty. (I)-(III) function by inactivating  
CC the DNA-binding activity of DP/E2F heterodimers. They are also  
CC used as research reagents, as positive controls in assays for  
CC identifying antagonists of DP-1/E2F dimerisation and as immunoassay  
CC agents. Also described is the use of sequences antisense to  
CC nucleic acids encoding (I)-(III) to control DP levels in cells,  
CC particularly by gene therapy. When formulated with cytotoxic  
CC or cytostatic agents, (I)-(III) enhance cell killing.  
XX SQ  
Sequence 37 AA;  
Query Match 100.0%; Score 31; DB 19; Length 37;  
Best Local Similarity 100.0%; Pred. No. 2.2; Mismatches 0; Indels 0; Gaps 0;  
Matches 7; Conservative 0;  
QY 1 ALNVLMA 7  
DB 10 ALNVLMA 16  
RESULT 11  
AAAY32163  
ID AAY32163 standard; Protein; 83 AA.  
XX AC AAY32163;  
XX DT 01-FEB-2000 (first entry)  
XX DE Soybean DP-1 protein fragment.  
XX KW DP-1; soybean; cell cycle regulatory protein;  
XX KW transcription factor; herbicide.  
XX OS Glycine max.  
XX FH Key Location/Qualifiers  
FT Misc-difference 10 /note= "encoded by GNC"  
FT FT Misc-difference 25 /note= "encoded by ANT"  
FT FT Misc-difference 26 /note= "encoded by GNC"  
FT FT

FT Misc-difference 34 /note= "encoded by GGN"  
FT FT Misc-difference 35 /note= "encoded by CNT"  
FT FT Misc-difference 49 /note= "encoded by NAN"  
FT FT Misc-difference 54 /note= "encoded by ANG"  
FT FT Misc-difference 55 /note= "encoded by NAT"  
FT FT Misc-difference 59 /note= "encoded by NAG"  
FT FT Misc-difference 63 /note= "encoded by NAT"  
FT FT Misc-difference 69 /note= "encoded by GNG"  
FT FT Misc-difference 71 /note= "encoded by NTC"  
FT FT Misc-difference 80 /note= "encoded by CNA"  
XX XX  
XX PN WO9953075-A2.  
XX PD 21-OCT-1999.  
XX PF 08-APR-1999; 99WO-US07638.  
XX PR 09-APR-1998; 98US-0081132.  
XX PA (DUPO ) DU PONT DE NEMOURS & CO E I.  
XX PI Klein TM, Morakinyo LO, Odell JT, Sakai H;  
XX DR WPI; 1999-633830/54.  
XX DR N-PSDB; AA234579.  
XX PT Plant-derived cell cycle regulatory proteins  
XX PS Claim 10; Page 41; 44pp; English.  
XX CC This sequence represents 42% of the middle region of soybean cell  
CC cycle regulatory protein DP-1, as deduced from an isolated  
CC cDNA clone (see AA234579). The invention relates to nucleic acid  
CC fragments (see AA234575-83) encoding plant CDC-16, DP-1, DP-2 and  
CC E2F cell cycle regulatory proteins (see AA234579-67). It also  
CC relates to the construction of a chimeric gene encoding all or a  
CC portion of the cell cycle regulatory protein, in sense or antisense  
CC orientation, where expression of the chimeric gene results in  
CC production of altered levels of the cell cycle regulatory protein in  
CC a transformed host cell. The nucleic acids and proteins may be  
CC used to facilitate studies of cell cycle regulation in plants,  
CC provide genetic tools to enhance cell growth in tissue culture,  
CC increase gene transfer efficiency and provide more stable  
CC transformations. The proteins may also provide targets to  
CC facilitate design and/or identification of cell cycle regulatory  
CC proteins that may be useful as herbicides.  
XX SQ  
Sequence 83 AA;  
Query Match 100.0%; Score 31; DB 20; Length 83;  
Best Local Similarity 100.0%; Pred. No. 5.5;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 ALNVLMA 7  
DB 13 ALNVLMA 19  
RESULT 12  
AAU72578  
ID AAU72578 standard; Peptide; 8 AA.  
XX XX  
XX AC AAU72578;



ID AAB60781 standard; protein; 56 AA.  
 AC AAB60781;  
 XX  
 DT 28-MAR-2001 (first entry)  
 DE Scorpion leuuropeptide I protein.  
 XX  
 XX Scorpion; toxin; K-channel; potassium; insect; pesticide.  
 XX  
 XX Hottentotta judaica.  
 XX  
 XX WO200078958-A2.  
 XX  
 XX 28-DEC-2000.  
 XX  
 XX 21-JUN-2000; 2000WO-US17049.  
 XX  
 XX 22-JUN-1999; 99US-0140227.  
 XX  
 XX (DUPO ) DU PONT DE NEMOURS & CO E I.  
 XX  
 XX Herrmann R, Lee J; Wong JF;  
 XX  
 XX WPI; 2001-071394/08.  
 XX  
 XX New polynucleotides encoding scorpion venom potassium-channel agonist  
 PT proteins for production e.g. of insect-tolerant transgenic plants for  
 PT controlling insect pest damage and parasitic worm infections -  
 XX  
 XX Claim 10; Fig 3; 50pp; English.  
 XX  
 XX The present invention relates to scorpion toxins. The invention may be  
 CC used for the creation of transgenic plants which express K-channel  
 CC modifiers, useful as a means for controlling insect pests by producing  
 CC insect-tolerant plants. In the prevention and/or treatment of insect  
 CC pest damage and parasitic worm infections in animals and humans, the  
 CC invention may also find use in creating specific new pesticides and  
 CC antihelmintic drugs that are also non-toxic to humans, pets and  
 CC livestock.  
 XX  
 XX Sequence 56 AA;  
 SQ  
 Query Match 87.1%; Score 27; DB 22; Length 56;  
 Best Local Similarity 71.4%; Pred. No. 33;  
 Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 ALNVLMA 7  
 DB 13 AMNVVMA 19  
 :||:|  
 RESULT 15  
 AAU72601  
 ID AAU72601 standard; Peptide; 35 AA.  
 XX  
 XX AAU72601;  
 XX  
 XX 26-FEB-2002 (first entry)  
 XX  
 XX DEF domain consensus sequence.  
 XX  
 XX Cell cycle protein; CCP; cell cycle regulation; herbicide;  
 KW plant growth regulator; plant development; abiotic stress; biotic stress;  
 KW nutrient deprivation; pathogen attack; crop yield; motif.  
 XX  
 XX Synthetic.  
 OS  
 XX WO200185946-A2.  
 XX  
 XX 15-NOV-2001.  
 XX  
 XX 14-MAY-2001; 2001WO-IB01307.  
 PF

XX 12-MAY-2000; 2000US-204045P.  
 XX  
 XX (CROP-) CROPDESIGN NV.  
 XX  
 XX Inze D, Boudolf V, De Veylder L, Acosta JAT, Magyar Z;  
 XX  
 XX WPI; 2002-062249/08.  
 XX  
 XX New cell cycle protein and nucleic acid molecule encoding it useful for  
 PT regulating cell cycle progression in plants and for identifying  
 PT modulators which are useful as herbicides or plant growth regulators -  
 XX  
 XX Disclosure; Page 25; 316pp; English.  
 XX  
 XX The invention relates to a novel cell cycle protein (CCP) and the  
 CC polynucleotides encoding them. CCP is useful for identifying a compound  
 CC which modulates the activity of the polypeptide and which binds to the  
 CC polypeptide and an anti-CCP antibody is useful for detecting the presence  
 CC of CCP in a sample. A CCP modulator is useful for modulating the cell  
 CC cycle or growth of a plant such as Arabidopsis thaliana, rice, wheat,  
 CC maize, tomato, alfalfa, oilseed rape, soybean, sunflower and canola.  
 CC CCP nucleic acid and polypeptide molecules are useful as modulating  
 CC agents in regulating cell cycle progression in plants. CCP is useful to  
 CC treat disorders characterised by insufficient or excessive production of  
 CC CCP protein or production of CCP protein forms which have decreased or  
 CC aberrant activity. Compounds that bind to or modulate the activity  
 CC of CCP polypeptide are useful as herbicides or plant growth regulators.  
 CC The polynucleotide is useful for modifying cell fate, plant development,  
 CC plant morphology, biochemistry and/or physiology, the length of the G1,  
 CC S, G2 and/or M phase of the cell cycle of a plant, initiation, promotion,  
 CC stimulation or enhancement of cell division, DNA replication, seed set,  
 CC seed size, seed development, tuber, fruit, leaf formation, shoot and root  
 CC initiation and/or development, nodule function, dwarfism in plants,  
 CC senescence, tolerance or resistance to stress. CCP, the polynucleotide  
 CC and the anti-CCP antibody are useful in agriculture to modulate the  
 CC protein levels or activity of a protein involved in the cell cycle due  
 CC to environmental conditions, including abiotic stress such as  
 CC cold, nutrient deprivation, heat, drought, salt stress, or biotic  
 CC stress such as pathogen attack, to modulate e.g. enhance crop yields,  
 CC and attenuate plant architecture, plant quality traits, plant  
 CC reproduction and seed development, endoreduplication in storage cells,  
 CC storage tissues and/or storage organs of plants or its parts. CCP is  
 CC useful as an immunogen to generate antibodies. CCP protein is useful to  
 CC screen for naturally occurring CCP substrates. The polynucleotide is  
 CC useful for expressing CCP protein, to detect CCP mRNA, or a genetic  
 CC lesion in a CCP gene and to modulate CCP activity. The present sequence  
 CC represents a motif which may be found in a CCP protein of the invention.  
 XX  
 XX Sequence 35 AA;  
 SQ  
 Query Match 83.9%; Score 26; DB 23; Length 35;  
 Best Local Similarity 85.7%; Pred. No. 33;  
 Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
 QY 1 ALNVLMA 7  
 DB 12 ALNVVMA 18  
 |||||  
 RESULT 16  
 AAB08901  
 ID AAB08901 standard; Protein; 41 AA.  
 XX  
 XX AAB08901;  
 XX  
 XX 30-AUG-2000 (first entry)  
 XX  
 XX Human secreted protein sequence encoded by gene 11 SEQ ID NO:58.  
 XX  
 XX Human; secreted protein; cytostatic; anti-proliferative; vulnery;  
 KW immunosuppressive; antibacterial; diagnosis; immune system; chemotaxis;  
 KW hyperproliferative disorder; infectious disease; tissue regeneration;  
 KW

KW screening; food additive; preservative; wound healing;  
 KW hyper-vascular disease.

OS Homo sapiens.

XX WO200017222-A1.

XX 30-MAR-2000.

XX 22-SEP-1999; 99WO-US22012.

XX 23-SEP-1998; 98US-0101546.

PR 02-OCT-1998; 98US-0102895.

XX (HUMA-) HUMAN GENOME SCI INC.

XX Ruben SM, Rosen CA, Duan RD, Shi Y, Lafleur DW, Young PE, Ni J;  
 PI Komatsoulis G, Endress GA, Soppet DR;

XX WPI; 2000-283538/24.

DR N-PSDB; AAA39062.

XX Human secreted proteins and coding sequences useful in diagnostic and  
 PT therapeutic methods for disorders such as immune system or  
 PT proliferative disorders, related to the proteins -

XX Claim 11; Page 358; 416pp; English.

XX The polynucleotide sequences given in AAA39052 to AAA39088 encode the  
 CC human secreted proteins given in AAB08891 to AAB08984. The human secreted  
 CC proteins can have activities based on the tissues and cells they are  
 CC expressed in. Examples of the activities are: cytostatic;  
 CC anti-proliferative; immunosuppressive; antibacterial; and vulnerary. The  
 CC secreted proteins and their related polynucleotide sequences are useful  
 CC for diagnostic and therapeutic methods useful for diagnosing and treating  
 CC disorders related to the secreted proteins. The proteins, and  
 CC polynucleotide sequences may be useful for treating disorders of the  
 CC immune system, hyperproliferative disorders, infectious diseases,  
 CC regeneration of tissues, for chemotaxis and for screening molecules that  
 CC bind to the proteins. The proteins or polynucleotide sequences may be  
 CC used as food additives or preservatives, to increase or decrease storage  
 CC capabilities, fat content, lipid, protein, carbohydrate, vitamins,  
 CC minerals, co-factors or other nutritional components. Agonists or  
 CC antagonists of the proteins may be used to prevent scar tissue growth  
 CC during wound healing, and hyper-vascular diseases. AAA39043 to AAA39051  
 CC and AAB08890 are sequences used in the exemplification of the present  
 CC invention.

XX Sequence 41 AA;

Query Match 83.9%; Score 26; DB 21; Length 41;

Best Local Similarity 85.7%; Pred. No. 40;

Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 ALNVLMA 7

Db 8 ALNVLFA 14

RESULT 17

AAB53284

ID AAB53284 standard; Protein; 90 AA.

XX AAB53284;

XX 09-MAR-2001 (first entry)

XX Human colon cancer antigen protein sequence SEQ ID NO:824.

XX Human; colon cancer; colon cancer antigen; diagnosis; detection;  
 KW identification; cytostatic; cardioactive; neuroprotective; vulnerary;  
 KW immunomodulatory; muscular; gynaecological; gastrointestinal;  
 KW nephrotropic; antiinfective; antibacterial; gene therapy; wound;

KW neural disorder; immune system disorder; muscular disorder;  
 KW reproductive disorder; gastrointestinal disorder; renal disorder;  
 KW infectious disease; cardiovascular disorder.

OS Homo sapiens.

XX WO200055351-A1.

XX 21-SEP-2000.

XX 08-MAR-2000; 2000WO-US05883.

XX 12-MAR-1999; 99US-0124270.

XX (HUMA-) HUMAN GENOME SCI INC.

XX Rosen CA, Ruben SM;

XX WPI; 2000-587534/55.

DR N-PSDB; AAC98041.

XX Colon cancer associated gene sequences, referred to as colon cancer

PT antigens, useful for the treatment, prevention, and diagnosis of colon

XX disorders such as colon cancer -

PS Claim 11; Page 1376; 2104pp; English.

XX AAC97991 to AAC98763 encode the human colon cancer associated proteins,  
 CC called human colon cancer antigens, given in AAB53234 to AAB54006. The  
 CC human colon cancer antigens can have cytostatic, cardioactive, muscular;  
 CC neuroprotective, immunomodulatory, gynaecological, gastrointestinal,  
 CC vulnerary, nephrotropic, antiinfective and antibacterial activities, and  
 CC can be used in gene therapy. The colon cancer antigen polynucleotides,  
 CC proteins and antibodies to the proteins are useful for the prevention,  
 CC treatment and diagnosis of colon disorders, such as colon cancer. The  
 CC polynucleotides may be used in diagnostics and research, such as for  
 CC chromosome identification, and as hybridisation probes. The proteins  
 CC may also be used to prevent diseases such as neural disorders, immune  
 CC system disorders, muscular disorders, reproductive disorders,  
 CC gastrointestinal disorders, wounds, renal disorders, infectious  
 CC diseases, and cardiovascular disorders. AAC98764 to AAC98772 and  
 CC AAB54007 represent sequences used in the exemplification of the present  
 CC invention.

XX Sequence 90 AA;

Query Match 83.9%; Score 26; DB 21; Length 90;

Best Local Similarity 83.3%; Pred. No. 1e+02;

Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ALNVLIM 6

Db 44 ALNVLIM 49

RESULT 18

AAO05442

ID AAO05442 standard; Protein; 59 AA.

XX AAO05442;

XX 06-NOV-2001 (first entry)

XX Human polypeptide SEQ ID NO 19334.

XX Human; cytokine; cell proliferation; cell differentiation; gene therapy;  
 KW vaccine; peptide therapy; stem cell growth factor; haematopoiesis;  
 KW tissue growth factor; immunomodulatory; cancer; leukaemia;  
 KW nervous system disorders; arthritis; inflammation.

OS Homo sapiens.

XX WO200164835-A2.

Tue Feb 17 11:55:55 2004

```

XX PD 07-SEP-2001.
XX PA (HYSE-) HYSEQ INC.
XX PI Tang YT, Liu C, Drmanac RT;
XX PI WPI; 2001-514838/56.
XX PR N-PSDB; AAI85373.
XX PR Isolated nucleic acids and polypeptides, useful for preventing
XX PR diagnosing and treating e.g. leukaemia, inflammation and immune
XX PR disorders -
XX PS Claim 20; SEQ ID NO 19334; 1399pp + Sequence Listing; English.
XX CC The invention relates to human polynucleotides (AAI79941-AAI93841) and
XX CC the encoded proteins (AAO00010-AAO13910) that exhibit activity elating to
XX CC cytokine, cell proliferation or cell differentiation or which may induce
XX CC production of other cytokines in other cell populations. The
XX CC polynucleotides and polypeptides are useful in gene therapy, vaccines or
XX CC peptide therapy. The polypeptides have various cytokine-like activities,
XX CC e.g. stem cell growth factor activity, haematopoiesis regulating
XX CC activity, tissue growth factor activity, immunomodulatory activity and
XX CC activin/inhibin activity and may be useful in the diagnosis and/or
XX CC treatment of cancer, leukaemia, nervous system disorders, arthritis and
XX CC inflammation.
XX CC Note: The sequence data for this patent did not form part of the printed
XX CC specification, but was obtained in electronic format directly from WIPO
XX CC at ftp.wipo.int/pub/published_pct_sequences.
XX SQ Sequence 59 AA;
    Query Match 80.6%; Score 25; DB 22; Length 59;
    Best Local Similarity 85.7%; Pred. No. 1.1e+02;
    Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
    QY 1 ALNVLMA 7
    Db 53 ALEVLMA 59
RESULT 19
AAI24066
ID AAM24066 standard; Protein; 73 AA.
XX AC AAM24066;
XX DE 12-OCT-2001 (first entry)
XX DE Human EST encoded protein SEQ ID NO: 1591.
XX KW Human; sheep; pig; cow; fruit fly; yeast; hamster; macaque; horse;
XX KW tomato; monkey; dog; sea urchin; expressed sequence tag; EST;
XX KW diagnostics; forensic test; gene mapping; genetic disorder;
XX KW biodiversity; gene therapy; nutrition.
XX OS Homo sapiens.
XX FN WO200154477-A2.
XX PD 02-AUG-2001.
XX PF 25-JAN-2001; 2001WO-US02687.
XX PR 25-JAN-2000; 2000US-0491404.
XX PR 17-JUL-2000; 2000US-0617746.
XX PR 03-AUG-2000; 2000US-0631451.

15-SEP-2000; 2000US-0663870.
XX PA (HYSE-) HYSEQ INC.
XX PI Tang YT, Liu C, Zhou P, Qian XB, Wang Z, Chen R, Asundi V;
XX PI Cao Y, Drmanac RA, Zhang J, Werhman T;
XX PR WPI; 2001-476164/51.
XX PR N-PSDB; AAI98725.
XX CC Isolated polypeptide for treatment of diseases, diagnostics, raising
XX CC antibodies and research use -
XX PS Claim 20; Page 1088; 1275pp; English.
XX CC The present invention provides the protein and coding sequences of novel
XX CC proteins from a variety of organisms, including human, dog, cat, horse,
XX CC cow, pig, hamster, monkey, macaque, yeast, bacteria, fruit fly, sea
XX CC urchin and tomato. These were derived from expressed sequence tags (ESTs)
XX CC from the organism of interest. They can be used in diagnostics,
XX CC forensics, gene mapping, identification of mutations, to assess
XX CC biodiversity and for nutritional purposes. The present sequence is a
XX CC protein of the invention.
XX SQ Sequence 73 AA;
    Query Match 80.6%; Score 25; DB 22; Length 73;
    Best Local Similarity 71.4%; Pred. No. 1.4e+02;
    Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;
    QY 1 ALNVLMA 7
    Db 2 SLNVLMA 8
RESULT 20
AAI75232
ID AAW75232 standard; Protein; 87 AA.
XX AC AAW75232;
XX DE 29-JAN-1999 (first entry)
XX DE Human secreted protein encoded by gene 20 clone HSDEG01.
XX KW Human; secreted protein; fusion protein; gene therapy; protein therapy;
XX KW diagnosis; tissue; cancer; tumour; neurodegenerative disorder; leukaemia;
XX KW developmental abnormality; foetal deficiency; blood; allergy; renal;
XX KW immune system; asthma; lymphocytic disease; brain; hepatic; lymphoma;
XX KW inflammation; ischaemic shock; Alzheimer's disease; restenosis; AIDS;
XX KW cognitive disorder; schizophrenia; prostate; obesity; osteoclast; thymus;
XX KW osteoporosis; arthritis; testis; lung; thyroiditis; thyroid; digestion;
XX KW endocrine; metabolism; regulation; malabsorption; gastritis; neoplasm.
XX OS Homo sapiens.
XX FN WO9840483-A2.
XX PD 17-SEP-1998.
XX PF 12-MAR-1998; 98WO-US04858.
XX PR 19-DEC-1997; 97US-0068368.
XX PR 14-MAR-1997; 97US-0040710.
XX PR 14-MAR-1997; 97US-0040762.
XX PR 30-MAY-1997; 97US-0048100.
XX PR 30-MAY-1997; 97US-0048189.

```



PR 30-MAY-1997; 97US-0048357.  
 PR 30-MAY-1997; 97US-0050934.  
 PR 06-JUN-1997; 97US-0048970.  
 PR 05-SEP-1997; 97US-0057765.  
 XX (HUMA-) HUMAN GENOME SCI INC.  
 XX  
 XX Ferrie AM, Fischer CL, Gentz RL, Greene JM, Kyaw H;  
 PI Li H, Li Y, Moore PA, Rosen CA, Ruben SM, Soppet DR;  
 PI Wei YF, Young PE, Zeng Z;  
 XX WPI: 1998-520811/44.  
 DR N-PSDB; AAV34322.  
 DR  
 XX Isolated human poly:nucleotide(s) encoding secretory peptide(s) -  
 PT used to develop products for the diagnosis and treatment of e.g.  
 PT inflammation, cancers, CNS disorders or immune system disorders  
 XX  
 XX Claim 1; Page 177; 201pp; English.  
 XX  
 CC This sequence represents a secreted human protein encoded by the gene  
 CC clone detailed in the descriptor line. The gene can be used to generate  
 CC fusion proteins by linking to the gene to a human immunoglobulin Fc  
 CC portion (e.g. AAV34277) for increasing the stability of the fused  
 CC protein as compared to the human protein only.  
 CC The invention relates to 28 novel genes and their fragments (nucleic  
 CC acid sequences: AAV34286-V34325; amino acid sequences AAW75196-W75235)  
 CC which are useful for preventing, treating or ameliorating medical  
 CC conditions e.g. by protein or gene therapy. Also, pathological  
 CC conditions can be diagnosed by determining the amount of the new  
 CC polypeptides in a sample or by determining the presence of mutations in  
 CC the new polynucleotides. Specific uses are described for each of the 28  
 CC polynucleotides, based on which tissues they are most highly expressed in  
 CC (see AAV34286 for described uses).  
 XX  
 SQ Sequence 87 AA;  
 Query Match 80.6%; Score 25; DB 19; Length 87;  
 Best Local Similarity 71.4%; Pred. No. 1.7e+02;  
 Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 ALNVLMA 7  
 Db :|||:|  
 2 SLNVLLA 8  
 RESULT 21  
 ID AAE26996 standard; Protein; 87 AA.  
 XX  
 AC AAE26996;  
 XX  
 XX 13-DEC-2002 (first entry)  
 XX  
 DE Human gene 20 encoded secreted protein HSDG01, SEQ ID NO:89.  
 XX  
 KW Human; immunodeficiency; X-linked agammaglobulinemia; septic shock;  
 KW autoimmune disorder; rheumatoid arthritis; multiple sclerosis; cancer;  
 KW Grave's disease; diabetes mellitus; haematopoietic disorder; stroke;  
 KW respiratory disorder; asthma; allergy; gastrointestinal disorder;  
 KW inflammatory bowel disease; neurodegenerative disorder; hepatitis;  
 KW Parkinson's disease; Alzheimer's disease; cardiovascular disorder;  
 KW atherosclerosis; myocarditis; renal disorder; fungicide; virucide;  
 KW hyperproliferative disorder; acute glomerulonephritis; tonsillitis;  
 KW respiratory disorder; rhinitis; sinusitis; neurological disease;  
 KW endocrine disorder; Addison's disease; reproductive system disorder;  
 KW endometriosis; vasotrophic; vulnary; cytostatic; nootropic; cardiant;  
 KW anti-HIV; tranquilliser; gout; antiparasitic.  
 XX  
 OS Homo sapiens.  
 XX  
 XX Key Location/Qualifiers  
 FH Peptide 1..21  
 FT

FT  
 FT Protein /label= Signal\_peptide  
 FT 22..87 /note= "Human mature secreted protein"  
 FT Misc-difference 11  
 FT /label= Unknown  
 FT /note= "Encoded by KGC"  
 FT Misc-difference 86  
 FT /label= Unknown  
 FT /note= "Encoded by AAN"  
 XX  
 XX US2002077287-A1.  
 XX  
 XX 20-JUN-2002.  
 XX  
 XX 11-MAY-2001; 2001US-0852659.  
 XX  
 XX 11-SEP-1998; 98US-0152060.  
 XX  
 XX (RUBE/) RUBEN S M.  
 XX (ROSE/) ROSEN C A.  
 XX (LIYY/) LI Y.  
 XX (ZENG/) ZENG Z.  
 XX (KYAW/) KYAW H.  
 XX (FISC/) FISCHER C L.  
 XX (LIHH/) LI H.  
 XX (SOPP/) SOPPET D R.  
 XX (GENT/) GENTZ R L.  
 XX (WEIY/) WEI Y.  
 PI Ruben SM, Rosen CA, Li Y, Zeng Z, Kyaw H, Fischer CL, Li H;  
 PI Soppet DR, Gentz RL, Wei Y, Moore PA, Young PE, Greene JM;  
 PI Ferrie AM;  
 XX WPI: 2002-598780/64.  
 DR N-PSDB; AAD44673.  
 DR  
 XX Novel human secreted polypeptides and polynucleotides for diagnosing,  
 PT preventing, treating immune, hyperproliferative, cardiovascular,  
 PT neurological, reproductive disorders and identifying modulators of  
 PT therapeutic use -  
 XX  
 PS Claim 11; Page 193; 209pp; English.  
 XX  
 XX AAD44636-AAD44676 represent cDNAs corresponding to 28 human secreted  
 CC protein genes, and AAE26959-AAE26999 represent the proteins they encode.  
 CC AAE27000-AAE27025 represent human secreted protein fragments or their  
 CC variants. The secreted proteins and genes are useful for preventing,  
 CC treating or ameliorating medical conditions, e.g., by protein or gene  
 CC therapy. Specific uses are described for each of the 28 genes, based  
 CC on the tissues in which they are most highly expressed and include  
 CC developing products for the diagnosis or treatment of immunodeficiencies,  
 CC e.g., X-linked agammaglobulinemia, B cell immunodeficiencies, severe  
 CC combined immunodeficiencies, autoimmune disorders e.g., systemic lupus  
 CC erythematosus, rheumatoid arthritis, multiple sclerosis, autoimmune  
 CC thyroiditis, autoimmune haemolytic anaemia, Goodpasture's syndrome,  
 CC Grave's disease, diabetes mellitus, dermatitis, inflammatory conditions  
 CC including septic shock, sepsis, reperfusion injury, inflammatory bowel  
 CC disease, Crohn's disease, haematopoietic disorders, respiratory  
 CC disorders e.g., asthma and allergy, gastrointestinal disorders e.g.,  
 CC inflammatory bowel disease), cancers e.g., gastric, ovarian, lung,  
 CC liver, bladder and breast), central nervous system (CNS) disorders e.g.,  
 CC ischaemic brain injury and/or stroke, neurodegenerative disorders e.g.,  
 CC Parkinson's disease and Alzheimer's disease, AIDS-related dementia and  
 CC prion disease, cardiovascular disorders e.g., myocarditis, arrhythmias,  
 CC atherosclerosis, inflammatory disorders e.g., hepatitis, gout, trauma,  
 CC pancreatitis, sarcoidosis and allogenic transplant rejection, blood-  
 CC related disorder (thrombosis, arterial thrombosis, atherosclerosis),  
 CC hyperproliferative disorders, respiratory disorders e.g. rhinitis,  
 CC sinusitis, tonsillitis, lung cancer, allergic disorders, pneumonitis,  
 CC renal disorders e.g. acute glomerulonephritis, neurological diseases,  
 CC liver disorders, endocrine disorders e.g., hyperthyroidism, Addison's  
 CC disease, hyperpituitarism, infectious diseases and reproductive system  
 CC disorders e.g. endometriosis. The present sequence represents a human

Tue Feb 17 11:55:55 2004

```

CC secreted protein of the invention.
XX
SQ Sequence 87 AA;
    Query Match 80.6%; Score 25; DB 23; Length 87;
    Best Local Similarity 71.4%; Pred. No. 1.7e+02;
    Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ALNVLMA 7
Db 2 SLNVLLA 8

RESULT 22
AAE27134
ID AAE27134 standard; Protein; 87 AA.
AC AAE27134;
XX
DT 13-DEC-2002 (first entry)
DE Human gene 20 encoded secreted protein HSDG01, SEQ ID NO:89.
XX
KW Human; secreted protein; autoimmune disease; hyperproliferative disorder;
KW rheumatoid arthritis; neoplasm; cerebrovascular disorder; angiogenesis;
KW cerebral ischaemia; cardiovascular disorder; nervous system disorder;
KW cardiac arrest; Alzheimer's disease; ocular disorder; wound healing;
KW infection; corneal infection; skin aging; food additive; preservative;
KW tissue regeneration; immunosuppressive; antiproliferative; cytostatic;
KW cardiant; vasotropic; cerebroprotective; nootropic; neuroprotective;
KW antibacterial; virucide; fungicide; ophthalmological; gene therapy;
KW
XX Homo sapiens.

XX Key Location/Qualifiers
FH Peptide 1..21 /label= signal_peptide
FT Protein 22..87
FT /note= "Mature human secreted protein"
FT Misc-difference 11
FT /note= "Encoded by KGC"
FT Misc-difference 86
FT /note= "Encoded by AAN"
XX
XX US2002076756-A1.
XX
XX 20-JUN-2002.
XX
XX 11-MAY-2001; 2001US-0853161.
XX
XX 02-FEB-2001; 2001US-265583P.
XX
XX (RUBE/) RUBEN S M.
XX (ROSE/) ROSEN C A.
XX (LIYI/) LI Y.
XX (ZENG/) ZENG Z.
XX (KYAW/) KYAW H.
XX (FISC/) FISCHER C L.
XX (LIH/) LI H.
XX (SOPP/) SOPPET D R.
XX (GENT/) GENTZ R L.
XX (WEIY/) WEI Y.
XX (MOOR/) MOORE P A.
XX (YOUN/) YOUNG P B.
XX (GREE/) GREENE J M.
XX (FERR/) FERRIE A M.
XX
XX Ruben SM, Rosen CA, Li Y, Zeng Z, Kyaw H, Fischer CL, Li H;
PI Soppet DR, Gentz RL, Wei Y, Moore PA, Young PE, Greene JM;
PI Ferrie AM;
XX WPI; 2002-574454/61.

```

---

```

DR N-PSDB; AAD44891.
XX New nucleic acid molecules encoding 28 human secreted proteins, useful
PT for diagnosing, preventing, treating or ameliorating medical conditions
PT and as food additives or preservatives -
XX Claim 11; Page 193-194; 209pp; English.
XX
XX AAD44854-AAD44984 represent cDNAs corresponding to 28 human secreted
CC protein genes, and AAE27097-AAE27137 represent the proteins they encode.
CC AAE27138-AAE27164 represent human secreted protein fragments. The genes
CC and their corresponding secreted proteins are useful for preventing,
CC treating or ameliorating medical conditions, e.g., by protein or gene
CC therapy. Secreted protein sequences of the invention are useful for the
CC diagnosis or treatment of disorders such as autoimmune diseases (e.g.
CC rheumatoid arthritis), hyperproliferative disorders (e.g. cerebral ischaemia,
CC the breast or liver), cerebrovascular disorders (e.g. cardiac arrest), nervous
CC angiogenesis), cardiovascular disorders (e.g. corneal infection). The
CC system disorders (e.g. Alzheimer's disease), infections caused by fungi,
CC bacteria and viruses and ocular disorders (e.g. corneal infection). The
CC polypeptides can also be used to aid wound healing and epithelial cell
CC proliferation, to prevent skin aging due to sunburn, to maintain organs
CC before transplantation, for supporting cell culture of primary tissues,
CC to regenerate tissues and in chemotaxis. They can also be used as food
CC additives or preservative to increase or decrease storage capabilities,
CC fat content, lipid, protein, carbohydrate, vitamins, minerals, cofactors
CC and other nutritional components. The present sequence represents a human
CC secreted protein of the invention.
XX
XX Sequence 87 AA;
SQ
    Query Match 80.6%; Score 25; DB 23; Length 87;
    Best Local Similarity 71.4%; Pred. No. 1.7e+02;
    Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ALNVLMA 7
Db 2 SLNVLLA 8

RESULT 23
ABU65007
ID ABU65007 standard; Protein; 87 AA.
XX
XX ABU65007;
XX
XX 15-MAY-2003 (first entry)
XX
XX Human secreted protein gene 20, protein #2.
XX
XX Secreted protein; immunodeficiency; multiple sclerosis;
XX severe combined immunodeficiency; autoimmune disorder; cancer;
XX rheumatoid arthritis; diabetes mellitus; haematopoietic disorder;
XX inflammatory condition; septic shock; inflammatory bowel disease;
XX Crohn's disease; respiratory disorder; asthma; allergy; stroke;
XX gastrointestinal disorder; central nervous system disorder;
XX ischaemic brain injury; neurodegenerative disorder; Parkinson's disease;
XX Alzheimer's disease; cardiovascular disorder; atherosclerosis;
XX blood-related disorder; thrombosis; atherosclerosis; renal disorder;
XX hyperproliferative disorder; acute glomerulonephritis; Addison's disease;
XX endocrine disorder; liver disease; reproductive system disorder;
XX endometriosis; infectious disease; pancreatic disorder; vaccine;
XX wound repair; angiogenesis; lymphatic disorder; hair loss; body weight;
XX body height; hair colour; human.
XX
XX Homo sapiens.
XX
XX US2002172994-A1.
XX
XX 21-NOV-2002.
XX
XX 11-MAY-2001; 2001US-0852797.
XX
XX

```



CC properties in comparison to the naturally glycosylated molecule.

```

XX SQ Sequence 25 AA;
Query Match 77.4%; Score 24; DB 19; Length 25;
Best Local Similarity 83.3%; Pred. No. 68;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 ALNVLM 6
DB 3 ALNVLL 8
|||||:

RESULT 25
AAW47300
ID AAW47300 standard; peptide; 25 AA.
XX
AC AAW47300;
XX
DT 01-JUN-1998 (first entry)
XX
DE Human IDS peptide fragment.
XX
KW Human; iduronate 2-sulphatase; IDS; treatment;
XX Hunter syndrome.
XX
OS Homo sapiens.
XX
PN US5728381-A.
XX
PD 17-MAR-1998.
XX
PF 07-JUN-1995; 95US-0484493.
XX
PR 17-DEC-1992; 92US-0991973.
PR 12-NOV-1991; 91US-0790362.
PR 28-NOV-1994; 94US-0345212.
PR 07-JUN-1995; 95US-0484493.
XX
PA (ANSO/) ANSON D S.
PA (BIEL/) BIELICKI J.
PA (CLEW/) CLEMENTS P R.
PA (HOPW/) HOPWOOD J J.
PA (MORR/) MORRIS C P.
PA (OCCH/) OCCHIODORO T.
PA (WILS/) WILSON P J.
XX
PI Anson DS, Bielicki J, Clements PR, Hopwood JJ, Morris CP;
PI Occhiodoro T, Wilson PJ;
XX
FS WPI; 1998-206530/18.
XX
DR Treatment of iduronate 2-sulphatase deficiency - comprises
PT administering recombinant iduronate 2-sulphatase
XX
XX Example 1; Column 11; 53pp; English.
XX
FS The present sequence is a human iduronate 2-sulphatase (IDS)
XX peptide fragment. IDS deficiency can be treated by administering a
XX recombinant human IDS that is more highly glycosylated than the
XX naturally occurring enzyme, useful in the treatment of Hunter
XX syndrome. The recombinant IDS may be administered in 0.5 microg/kg
XX to 20 mg/kg doses. The administration route is oral, intravenous,
XX intraperitoneal, intramuscular, subcutaneous or intranasal. The
XX recombinant IDS has better uptake properties and/or a longer
XX half-life in vivo, and is thus more efficient than naturally
XX glycosylated IDS.
XX
SQ Sequence 25 AA;
Query Match 77.4%; Score 24; DB 19; Length 25;
Best Local Similarity 83.3%; Pred. No. 68;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

```

```

QY 1 ALNVLM 6
DB 3 ALNVLL 8
|||||:

RESULT 26
AAAY23983
ID AAY23983 standard; Peptide; 25 AA.
XX
AC AAY23983;
XX
DT 21-SEP-1999 (first entry)
XX
DE Amino terminal sequence of iduronate 2-sulphatase protein.
XX
KW Human; iduronate 2-sulphatase; IDS; IDS deficiency; Hunters syndrome.
XX
OS Homo sapiens.
XX
PN US5932211-A.
XX
PD 03-AUG-1999.
XX
PF 28-NOV-1994; 94US-0345212.
XX
PR 17-DEC-1992; 92US-0991973.
PR 12-NOV-1991; 91US-0790362.
PR 28-NOV-1994; 94US-0345212.
XX
PA (WOME-) WOMEN'S & CHILDREN'S HOSPITAL.
XX
PI Anson DS, Bielicki J, Clements PR, Hopwood JJ, Morris CP;
PI Occhiodoro T, Wilson PJ;
XX
DR WPI; 1999-443569/37.
XX
PT Recombinant iduronate 2-sulphatase useful for treating Hunters
PT syndrome
XX
XX Example 1; Column 11; 54pp; English.
XX
FS The present sequence is derived from human iduronate 2-sulphatase (IDS)
XX protein, and is used to design a probe for the isolation of cDNA
XX encoding IDS. The specification describes recombinant human IDS produced
XX in eukaryotic cells that has enzymatic activity and is more highly
XX glycosylated than native IDS from the human liver, placenta or
XX kidney. The protein is useful for treating IDS deficiency,
XX e.g. Hunters syndrome.
XX
SQ Sequence 25 AA;
Query Match 77.4%; Score 24; DB 20; Length 25;
Best Local Similarity 83.3%; Pred. No. 68;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 ALNVLM 6
DB 3 ALNVLL 8
|||||:

RESULT 27
AAB51178
ID AAB51178 standard; Peptide; 25 AA.
XX
AC AAB51178;
XX
DT 21-MAR-2001 (first entry)
XX
DE Human iduronate 2-sulphatase (IDS) peptide sequence SEQ ID NO:5.
XX
KW Human; iduronate 2-sulphatase; IDS; glycosylated; gene therapy;
KW mucopolysaccharidosis inhibitor; IDS deficiency disorder;

```

KW Hunter syndrome; mucopolysaccharidosis type II.  
 XX Homo sapiens.  
 XX US6153188-A.  
 XX 28-NOV-2000.  
 XX 12-FEB-1999; 99US-0249003.  
 XX 17-DEC-1992; 92US-0991973.  
 PR 28-NOV-1994; 94US-0345212.  
 PR 12-NOV-1991; 91US-0790362.  
 XX (WOME-) WOMEN'S & CHILDREN'S HOSPITAL.  
 XX Wilson PJ, Anson DS, Occhiodoro T, Bielicki J, Clements PR;  
 PI Hopwood JJ, Morris CP;  
 XX WPI; 2001-060076/07.  
 XX New highly glycosylated recombinant human iduronate 2-sulfatase (IDS)  
 PT useful for diagnosing or treating subjects suspected of having or  
 PT suffering from IDS deficiency disorders, e.g. Hunter syndrome  
 PT (mucopolysaccharidosis-II) -  
 XX Example 1; Column 11; 53pp; English.  
 XX The present invention describes a recombinant human iduronate 2-sulfatase  
 CC (IDS). The recombinant IDS is more highly glycosylated than the naturally  
 CC occurring enzyme isolated from human tissue. The recombinant human IDS  
 CC can be produced in Chinese Hamster Ovary (CHO) cells or in a human cell.  
 CC The recombinant IDS comprises a fusion protein. It is a  
 CC mucopolysaccharidosis inhibitor and can be used in gene therapy.  
 CC The recombinant IDS is useful in treating and diagnosing subjects  
 CC suffering from or suspected of having IDS deficiency disorders, e.g.  
 CC Hunter syndrome (mucopolysaccharidosis type II). The present sequence  
 CC represents a human IDS peptide which is used to produce an  
 CC oligonucleotide probe used in an example from the present invention.  
 XX SQ Sequence 25 AA;  
 Query Match 77.4%; Score 24; DB 22; Length 25;  
 Best Local Similarity 83.3%; Pred. No. 68;  
 Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 ALNVLM 6  
 DB 3 ALNVLL 8  
 RESULT 28  
 ABG49596  
 ID ABG49596 standard; Peptide; 27 AA.  
 XX ABG49596;  
 XX 25-FEB-2003 (first entry)  
 XX Human liver peptide, SEQ ID No 28244.  
 XX Human; liver; cirrhosis; hyperlipoproteinaemia; hyperlipidaemia;  
 KW hypercholesterolaemia; coronary heart disease.  
 XX Homo sapiens.  
 XX WO200157273-A2.  
 XX 09-AUG-2001.  
 XX 30-JAN-2001; 2001WO-US00664.  
 XX 04-FEB-2000; 2000US-0180312.  
 PR 26-MAY-2000; 2000US-0207456.  
 PR 30-JUN-2000; 2000US-0608408.  
 PR 03-AUG-2000; 2000US-0632366.  
 PR 21-SEP-2000; 2000US-0234687.  
 PR 27-SEP-2000; 2000US-0236359.  
 PR 04-OCT-2000; 2000GB-0024263.  
 XX (MOLE-) MOLECULAR DYNAMICS INC.  
 XX Penn SG, Hanzel DK, Chen W, Rank DR;  
 WPI; 2001-488898/53.  
 XX Human genome-derived single exon nucleic acid probes useful for  
 PT analysing gene expression in human adult liver -  
 XX Claim 27; SEQ ID No 28244; 658pp; English.  
 XX The invention relates to a single exon nucleic acid probe (SENP) (I) for  
 CC measuring human gene expression in a sample derived from human adult  
 CC liver, comprising one of 13109 defined nucleotide sequences given in the  
 CC specification (or complements/ fragments). The probe hybridises at high  
 CC stringency to a nucleic acid molecule expressed in the human adult  
 CC liver. (I) may be used for predicting, measuring and displaying gene  
 CC expression in samples derived from human adult liver. The genes  
 CC identified may be involved in genetic liver diseases such as cirrhosis,  
 CC hyperlipoproteinaemia, hyperlipidaemia and hypercholesterolaemia which  
 CC is associated with coronary heart disease. ABG47348-ABG59930 represent  
 CC human liver single exon encoded peptides of the invention.  
 CC Note: The sequence information for this patent does not appear in the  
 CC printed specification but was obtained in electronic format directly  
 CC from WIPO at ftp.wipo.int/pub/published\_pct\_sequences.  
 XX SQ Sequence 27 AA;  
 Query Match 77.4%; Score 24; DB 22; Length 27;  
 Best Local Similarity 83.3%; Pred. No. 75;  
 Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;  
 QY 2 LNVLM 7  
 DB 4 LNVLL 9  
 RESULT 29  
 ABB29588  
 ID ABB29588 standard; Peptide; 27 AA.  
 XX ABB29588;  
 XX 01-FEB-2002 (first entry)  
 XX Peptide #2239 encoded by breast cell single exon nucleic acid probe.  
 DE Human; microarray; single exon probe; gene expression; breast;  
 KW disease; cancer.  
 XX Homo sapiens.  
 XX WO200157271-A2.  
 XX 09-AUG-2001.  
 XX 30-JAN-2001; 2001WO-US00662.  
 XX 04-FEB-2000; 2000US-0180312.  
 PR 26-MAY-2000; 2000US-0207456.  
 PR 30-JUN-2000; 2000US-0608408.  
 PR 03-AUG-2000; 2000US-0632366.  
 PR 21-SEP-2000; 2000US-0234687.  
 PR 27-SEP-2000; 2000US-0236359.  
 PR 04-OCT-2000; 2000GB-0024263.  
 XX

```
PA (MOLE-) MOLECULAR DYNAMICS INC.
XX Penn SG, Hanzel DK, Chen W, Rank DR;
PI WPI; 2001-496933/54.
XX Human genome-derived single exon nucleic acid probes useful for
XX analyzing gene expression in human foetal liver -
XX Claim 27; SEQ ID NO 12556; 327pp + sequence listing; English.
XX The invention relates to a spatially-addressable set of single exon
XX nucleic acid probes for measuring gene expression in a sample derived
XX from human breast and BT 474 cells. The method involves contacting
XX the probes with a collection of detectably labelled nucleic acids
XX derived from mRNA of human breast, and then measuring the label
XX bound to each probe of the microarray. The probes are useful for
XX verifying the expression of regions of genomic DNA predicted to
XX encode proteins. They are useful for gene discovery, and for
XX determining predisposition and/or prognosing breast disease. Gene
XX expression analysis is useful for assessing the toxicity of chemical
XX agents on cells. The microarray of this invention presents a far greater
XX diversity of probes for measuring gene expression, with far less bias
XX than expressed sequence tag microarrays. The method is suitable for
XX rapid production of functional information from genomic sequence. The
XX present sequence is a peptide encoded by a single exon nucleic acid
XX probe of the invention.
XX Note: The sequence data for this patent did not form part of the
XX printed specification, but was obtained in electronic format directly
XX from WIPO at ftp.wipo.int/pub/published_pct_sequences.
XX SQ Sequence 27 AA;
XX Query Match 77.4%; Score 24; DB 22; Length 27;
XX Best Local Similarity 83.3%; Pred. No. 75;
XX Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
QY 2 LNVLMA 7
Db ||:||||
4 LNLMA 9
RESULT 30
ABB34768
ID ABB34768 standard; Peptide; 27 AA.
XX AC ABB34768;
XX DT 04-FEB-2002 (first entry)
XX DE Peptide #2274 encoded by human foetal liver single exon probe.
XX KW Human; foetal liver; gene expression; single exon nucleic acid probe.
XX OS Homo sapiens.
XX PN WO200157277-A2.
XX PD 09-AUG-2001.
XX PF 30-JAN-2001; 2001WO-US00669.
XX 04-FEB-2000; 2000US-0180312.
XX 26-MAY-2000; 2000US-0207456.
XX 30-JUN-2000; 2000US-0608408.
XX 03-AUG-2000; 2000US-0632366.
XX 21-SEP-2000; 2000US-0234687.
XX 27-SEP-2000; 2000US-0236359.
XX 04-OCT-2000; 2000GB-0024263.
XX (MOLE-) MOLECULAR DYNAMICS INC.
XX PA Penn SG, Hanzel DK, Chen W, Rank DR;
XX PI WPI; 2001-498899/53.
XX DR Single exon nucleic acid probes for analyzing gene expression in human
XX PT hearts -
XX PT hearts -
XX XX Claim 15; SEQ ID No 21953; 530pp; English.
XX The present invention relates to single exon nucleic acid probes for
XX measuring human gene expression in a sample derived from human heart (see
```

```
PI Penn SG, Hanzel DK, Chen W, Rank DR;
XX WPI; 2001-483447/52.
XX Human genome-derived single exon nucleic acid probes useful for
XX analyzing gene expression in human foetal liver -
XX Claim 27; SEQ ID NO 27403; 639pp + sequence listing; English.
XX The invention relates to a single exon nucleic acid probe for
XX measuring human gene expression in a sample derived from human foetal
XX liver. The single exon nucleic acid probes may be used for predicting,
XX measuring and displaying gene expression in samples derived from human
XX foetal liver. The present sequence is a peptide encoded by a single exon
XX nucleic acid probe of the invention.
XX Note: The sequence data for this patent did not form part of the
XX printed specification, but was obtained in electronic format directly
XX from WIPO at ftp.wipo.int/pub/published_pct_sequences.
XX SQ Sequence 27 AA;
XX Query Match 77.4%; Score 24; DB 22; Length 27;
XX Best Local Similarity 83.3%; Pred. No. 75;
XX Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
QY 2 LNVLMA 7
Db ||:||||
4 LNLMA 9
RESULT 31
ABB20183
ID ABB20183 standard; Protein; 27 AA.
XX AC ABB20183;
XX DT 23-JAN-2002 (first entry)
XX DE Protein #2182 encoded by probe for measuring heart cell gene expression.
XX KW Human; gene expression; heart; microarray; vascular system;
XX KW cardiovascular disease; hypertension; cardiac arrhythmia;
XX KW congenital heart disease.
XX OS Homo sapiens.
XX PN WO200157274-A2.
XX PD 09-AUG-2001.
XX PF 30-JAN-2001; 2001WO-US00666.
XX 04-FEB-2000; 2000US-0180312.
XX 26-MAY-2000; 2000US-0207456.
XX 30-JUN-2000; 2000US-0608408.
XX 03-AUG-2000; 2000US-0632366.
XX 21-SEP-2000; 2000US-0234687.
XX 27-SEP-2000; 2000US-0236359.
XX 04-OCT-2000; 2000GB-0024263.
XX (MOLE-) MOLECULAR DYNAMICS INC.
XX PA Penn SG, Hanzel DK, Chen W, Rank DR;
XX PI WPI; 2001-498899/53.
XX DR Single exon nucleic acid probes for analyzing gene expression in human
XX PT hearts -
XX PT hearts -
XX XX Claim 15; SEQ ID No 21953; 530pp; English.
XX The present invention relates to single exon nucleic acid probes for
XX measuring human gene expression in a sample derived from human heart (see
```

CC ABA21535-ABM411305). The present sequence is a protein encoded by one such  
CC probe. The probes may be used for predicting, measuring and displaying  
CC gene expression in samples derived from the human heart via microarrays.  
CC By measuring gene expression, the probes are useful for predicting,  
CC diagnosing, grading, staging, monitoring and prognosing diseases of the  
CC human heart and vascular system e.g. cardiovascular disease,  
CC hypertension, cardiac arrhythmias and congenital heart disease.  
CC Note: The sequence data for this patent did not form part of the printed  
CC specification, but was obtained in electronic format directly from WIPO  
CC at ftp.wipo.int/pub/published\_pct\_sequences.

XX SQ Sequence 27 AA;

Query Match 77.4%; Score 24; DB 22; Length 27;  
Best Local Similarity 83.3%; Pred. No. 75;  
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 2 LNVLMMA 7  
Db 4 LNLMLMA 9

RESULT 32

AA055570  
ID AA055570 standard; Protein; 27 AA.

XX AC AA055570;

DT 05-NOV-2001 (first entry)

XX Human brain expressed single exon probe encoded protein SEQ ID NO: 27675.

DE Human; brain expressed exon; gene expression analysis; probe;  
XX microarray; Alzheimer's disease; multiple sclerosis; schizophrenia;  
KW epilepsy; cancer.

XX OS Homo sapiens.

XX WO200157275-A2.

XX PD 09-AUG-2001.

XX PF 30-JAN-2001; 2001WO-US00667.

XX PR 04-FEB-2000; 2000US-0180312.

XX PR 26-MAY-2000; 2000US-0207456.

XX PR 30-JUN-2000; 2000US-0608408.

XX PR 03-AUG-2000; 2000US-0632366.

XX PR 21-SEP-2000; 2000US-0234687.

XX PR 27-SEP-2000; 2000US-0236359.

XX PR 04-OCT-2000; 2000GB-0024263.

XX PA (MOLE-) MOLECULAR DYNAMICS INC.

XX PI Penn SG, Hanzel DK, Chen W, Rank DR;

XX WPI; 2001-483446/52.

XX Single exon nucleic acid probes for analyzing gene expression in human  
PT brains -

XX Example 4; SEQ ID NO: 27675; 650pp + Sequence Listing; English.

XX The present invention provides a number of single exon nucleic acid  
CC probes which are derived from genomic sequences expressed in the human  
CC brain. They can be used to measure gene expression in brain cell samples,  
CC which may enable the diagnosis and improved treatment of nervous system  
CC diseases such as Alzheimer's disease, multiple sclerosis, schizophrenia,  
CC epilepsy and cancers. The present sequence is a protein encoded by one of  
CC the probes of the invention.

XX SQ Sequence 27 AA;

Query Match 77.4%; Score 24; DB 22; Length 27;  
Best Local Similarity 83.3%; Pred. No. 75;  
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 2 LNVLMMA 7  
Db 4 LNLMLMA 9

RESULT 33

AA067954  
ID AA067954 standard; Protein; 27 AA.

XX AC AA067954;

DT 06-NOV-2001 (first entry)

XX Human bone marrow expressed probe encoded protein SEQ ID NO: 28260.

DE Human; bone marrow expressed exon; gene expression analysis; probe;  
XX microarray; cancer; leukaemia; lymphoma; myeloma.

XX OS Homo sapiens.

XX WO200157276-A2.

XX PD 09-AUG-2001.

XX PF 30-JAN-2001; 2001WO-US00668.

XX PR 04-FEB-2000; 2000US-0180312.

XX PR 26-MAY-2000; 2000US-0207456.

XX PR 30-JUN-2000; 2000US-0608408.

XX PR 03-AUG-2000; 2000US-0632366.

XX PR 21-SEP-2000; 2000US-0234687.

XX PR 27-SEP-2000; 2000US-0236359.

XX PR 04-OCT-2000; 2000GB-0024263.

XX PA (MOLE-) MOLECULAR DYNAMICS INC.

XX PI Penn SG, Hanzel DK, Chen W, Rank DR;

XX WPI; 2001-488900/53.

XX Human genome-derived single exon nucleic acid probes useful for  
PT analyzing gene expression in human bone marrow -

XX Example 4; SEQ ID NO: 28260; 659pp + Sequence Listing; English.

XX The present invention provides a number of single exon nucleic acid  
CC probes which are derived from genomic sequences expressed in the human  
CC bone marrow. They can be used to measure gene expression in bone marrow  
CC samples, which may enable the improved diagnosis and treatment of cancers  
CC such as lymphoma, leukaemia and myeloma. The present sequence is a  
CC protein encoded by one of the probes of the invention.

XX SQ Sequence 27 AA;

Query Match 77.4%; Score 24; DB 22; Length 27;  
Best Local Similarity 83.3%; Pred. No. 75;  
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 2 LNVLMMA 7  
Db 4 LNLMLMA 9

RESULT 34

AA015772  
ID AA015772 standard; Protein; 27 AA.

XX AC AA015772;

XX

DT 12-OCT-2001 (first entry)  
XX Peptide #2206 encoded by probe for measuring cervical gene expression.  
DE Probe; human; microarray; gene expression; cervical epithelial cell;  
XX Cervical cancer.  
XX Homo sapiens.  
XX WO200157278-A2.  
XX 09-AUG-2001.  
XX 30-JAN-2001; 2001WO-US00670.  
XX 04-FEB-2000; 2000US-0180312.  
XX 26-MAY-2000; 2000US-0207456.  
XX 30-JUN-2000; 2000US-0608408.  
XX 03-AUG-2000; 2000US-0632366.  
XX 21-SEP-2000; 2000US-0234687.  
XX 27-SEP-2000; 2000US-0236359.  
XX 04-OCT-2000; 2000GB-0024263.  
XX (MOLE-) MOLECULAR DYNAMICS INC.  
XX Penn SG, Hanzel DK, Chen W, Rank DR;  
XX WPI; 2001-488901/53.  
XX Human genome-derived single exon nucleic acid probes useful for  
XX analyzing gene expression in human cervical epithelial cells -  
XX Claim 27; SEQ ID No 20598; 487pp; English.  
XX The present invention relates to human single exon nucleic acid probes  
XX (SENPs; see AAI10068-AAI28459). The present sequence is a peptide encoded  
XX by one such probe. The SENPs are derived from human HeLa cells. The SENPs  
XX can be used to produce a single exon microarray, which can be used for  
XX measuring human gene expression in a sample derived from human cervical  
XX epithelial cells. By measuring gene expression, the probes are therefore  
XX useful in grading and/or staging of diseases of the cervix, notably  
XX cervical cancer.  
XX Note: The sequence data for this patent did not form part of the printed  
XX specification, but was obtained in electronic format directly from WIPO  
XX at ftp.wipo.int/pub/published\_pct\_sequences.  
XX SQ Sequence 27 AA;  
Query Match 77.4%; Score 24; DB 22; Length 27;  
Best Local Similarity 83.3%; Pred. No. 75;  
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;  
QY 2 LNVLMVA 7  
Db 4 LNLMLA 9  
RESULT 35  
AAM28281  
ID AAM28281 standard; Protein; 27 AA.  
XX AC AAM28281;  
XX 17-OCT-2001 (first entry)  
XX Peptide #2318 encoded by probe for measuring placental gene expression.  
DE Probe; microarray; human; placenta; antenatal diagnosis;  
XX genetic disorder.  
XX Homo sapiens.  
XX WO200157272-A2.  
XX

XX 09-AUG-2001.  
XX 30-JAN-2001; 2001WO-US00663.  
XX 04-FEB-2000; 2000US-0180312.  
XX 26-MAY-2000; 2000US-0207456.  
XX 30-JUN-2000; 2000US-0608408.  
XX 03-AUG-2000; 2000US-0632366.  
XX 21-SEP-2000; 2000US-0234687.  
XX 27-SEP-2000; 2000US-0236359.  
XX 04-OCT-2000; 2000GB-0024263.  
XX (MOLE-) MOLECULAR DYNAMICS INC.  
XX Penn SG, Hanzel DK, Chen W, Rank DR;  
XX WPI; 2001-488977/53.  
XX Human genome-derived single exon nucleic acid probes useful for  
XX analyzing gene expression in human placenta -  
XX Claim 27; SEQ ID No 28550; 654pp; English.  
XX The present invention relates to single exon nucleic acid probes (SENPs;  
XX see AAI31315-AAI57546). The present sequence is a peptide encoded by one  
XX such probe. The probes are useful for producing a microarray for  
XX predicting, measuring and displaying gene expression in samples derived  
XX from human placenta. The probes are useful for antenatal diagnosis of  
XX human genetic disorders.  
XX SQ Sequence 27 AA;  
Query Match 77.4%; Score 24; DB 22; Length 27;  
Best Local Similarity 83.3%; Pred. No. 75;  
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;  
QY 2 LNVLMVA 7  
Db 4 LNLMLA 9  
RESULT 36  
AAM03506  
ID AAM03506 standard; Protein; 27 AA.  
XX AC AAM03506;  
XX 09-OCT-2001 (first entry)  
XX Peptide #2188 encoded by probe for measuring breast gene expression.  
DE Probe; human; breast disease; breast cancer; development disorder;  
XX inflammatory disease; proliferative breast disease; non-carcinoma tumour.  
XX Homo sapiens.  
XX WO200157270-A2.  
XX 09-AUG-2001.  
XX 29-JAN-2001; 2001WO-US00661.  
XX 04-FEB-2000; 2000US-0180312.  
XX 26-MAY-2000; 2000US-0207456.  
XX 30-JUN-2000; 2000US-0608408.  
XX 03-AUG-2000; 2000US-0632366.  
XX 21-SEP-2000; 2000US-0234687.  
XX 27-SEP-2000; 2000US-0236359.  
XX 04-OCT-2000; 2000GB-0024263.  
XX (MOLE-) MOLECULAR DYNAMICS INC.  
XX



PI Penn SG, Hanzel DK, Chen W, Rank DR;  
 XX WPI; 2001-476286/51.  
 XX Novel single exon nucleic acid probe used to measuring gene expression  
 PT in a human breast -  
 PT  
 XX Claim 27; SEQ ID No 12246; 322pp; English.  
 XX The present invention relates to novel single exon nucleic acid probes  
 CC (see AA100010-AA110067). The present sequence is a peptide encoded by one  
 CC such probe. The probes are useful for measuring human gene expression in a  
 CC human breast sample, where the probe hybridizes at high stringency to a  
 CC nucleic acid expressed in the human breast. The probes are useful for  
 CC predicting, diagnosing, grading, staging, monitoring and prognosing  
 CC diseases of the human breast, particularly those diseases with polygenic  
 CC aetiology. The diseases include: breast cancer, disorders of development,  
 CC inflammatory diseases of the breast, fibrocystic changes, proliferative  
 CC breast disease and non-carcinoma tumours.  
 CC Note: The sequence data for this patent did not form part of the printed  
 CC specification, but was obtained in electronic format directly from WIPO  
 CC at ftp.wipo.int/pub/published\_pct\_sequences.  
 XX  
 SQ Sequence 27 AA;  
 Query Match 77.4%; Score 24; DB 22; Length 27;  
 Best Local Similarity 83.3%; Pred. No. 75;  
 Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;  
 QY 2 LNVLMA 7  
 Db 4 LNLMLA 9  
 ||:||||  
 ||:||||  
 RESULT 37  
 ABG37489  
 ID ABG37489 standard; Peptide; 27 AA.  
 AC ABG37489;  
 XX  
 XX 19-AUG-2002 (first entry)  
 DT Human peptide encoded by genome-derived single exon probe SEQ ID 27154.  
 DE  
 DE Human; single exon probe; asthma; lung cancer; COPD; ILD;  
 KW chronic obstructive pulmonary disease; interstitial lung disease;  
 KW familial idiopathic pulmonary fibrosis; neurofibromatosis;  
 KW tuberous sclerosis; Gaucher's disease; Niemann-Pick disease;  
 KW Hermansky-Pudlak syndrome; sarcoidosis; pulmonary haemosiderosis;  
 KW pulmonary histiocytosis; lymphangioleiomyomatosis; Karagener syndrome;  
 KW pulmonary alveolar proteinosis; fibrocystic pulmonary dysplasia;  
 KW primary ciliary dyskinesia; pulmonary hypertension;  
 KW hyaline membrane disease.  
 XX Homo sapiens.  
 XX WO200186003-A2.  
 PN 15-NOV-2001.  
 XX  
 XX 30-JAN-2001; 2001WO-US00665.  
 XX  
 XX 04-FEB-2000; 2000US-180312P.  
 PR 26-MAY-2000; 2000US-207458P.  
 PR 30-JUN-2000; 2000US-0608408.  
 PR 03-AUG-2000; 2000US-0632366.  
 PR 21-SEP-2000; 2000US-234687P.  
 PR 27-SEP-2000; 2000US-236359P.  
 PR 04-OCT-2000; 2000GB-0024263.  
 XX (MOLE-) MOLECULAR DYNAMICS INC.  
 PA Penn SG, Hanzel DK, Chen W, Rank DR;  
 XX

XX WPI; 2002-114183/15.  
 XX Spatially-addressable set of single exon nucleic acid probes, used to  
 PT measure gene expression in human lung samples -  
 XX  
 XX Claim 27; SEQ ID No 27154; 634pp; English.  
 XX The invention relates to a spatially-addressable set of single exon  
 CC nucleic acid probes for measuring gene expression in a sample derived  
 CC from human lung comprising single exon nucleic acid probes having one of  
 CC 12614 nucleic acid sequences mentioned in the specification, or their  
 CC complements or the 12387 open reading frames derived from the 12614  
 CC probes. Also included are a microarray comprising the novel set of  
 CC probes; the novel set of probes which hybridise at high stringency to a  
 CC nucleic acid expressed in the human lung; measuring gene expression in a  
 CC sample derived from human lung, comprising (a) contacting the array with  
 CC a collection of detectably labeled nucleic acids derived from human lung  
 CC mRNA, and (b) measuring the label detectably bound to each probe of  
 CC the array; identifying exons in a eukaryotic genome, comprising  
 CC (a) algorithmically predicting at least one exon from genomic sequences  
 CC of the eukaryote; and (b) detecting specific hybridisation of detectably  
 CC labeled nucleic acids from eukaryote lung mRNA, to a single exon probe,  
 CC having a fragment identical to the predicted exon, the probe is included  
 CC in the above mentioned microarray; assigning exons to a single gene,  
 CC comprising (a) identifying exons from genomic sequence by the method  
 CC above and (b) measuring the expression of each of the exons in several  
 CC tissues and/or cell types using hybridisation to a single exon  
 CC microarrays having a probe with the exon, where a common pattern of  
 CC expression of the exons in the tissues and/or cell types indicates that  
 CC the exons should be assigned to a single gene; a peptide comprising one  
 CC of 12011 sequences, mentioned in the specification, or encoded by the  
 CC probes/open reading frames (ORF). The probes are used for gene  
 CC expression analysis, and for identifying exons in a gene, particularly  
 CC using human lung derived mRNA and for the study of lung diseases  
 CC such as asthma, lung cancer, chronic obstructive pulmonary disease  
 CC (COPD), interstitial lung disease (ILD), familial idiopathic pulmonary  
 CC fibrosis, neurofibromatosis, tuberous sclerosis, Gaucher's disease,  
 CC Niemann-Pick disease, Hermansky-Pudlak syndrome, sarcoidosis, pulmonary  
 CC haemosiderosis, pulmonary histiocytosis, lymphangioleiomyomatosis,  
 CC pulmonary alveolar proteinosis, Karagener syndrome, fibrocystic  
 CC pulmonary dysplasia, primary ciliary dyskinesia, pulmonary hypertension  
 CC and hyaline membrane disease. The present sequence is a peptide/protein  
 CC encoded by a single exon probe of the invention.  
 CC Note: The sequence data for this patent did not form part  
 CC of the printed specification, but was obtained in electronic  
 CC format directly from WIPO at  
 CC ftp.wipo.int/pub/published\_pct\_sequences.  
 XX  
 SQ Sequence 27 AA;  
 Query Match 77.4%; Score 24; DB 23; Length 27;  
 Best Local Similarity 83.3%; Pred. No. 75;  
 Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;  
 QY 2 LNVLMA 7  
 Db 4 LNLMLA 9  
 ||:||||  
 ||:||||  
 RESULT 38  
 AAG77060  
 ID AAG77060 standard; Protein; 32 AA.  
 XX  
 AC AAG77060;  
 XX  
 DT 03-SEP-2001 (first entry)  
 XX Human colon cancer antigen protein SEQ ID NO:7824.  
 DE  
 DE Human; colon cancer; colon cancer antigen; diagnosis; detection;  
 KW colorectal carcinoma.  
 XX

Tue Feb 17 11:55:55 2004

us-09-900-147-7-rag

OS Homo sapiens.  
XX WO200122920-A2.  
XX PD 05-APR-2001.  
XX PF 28-SEP-2000; 2000WO-US26524.  
XX PR 29-SEP-1999; 99US-0157137.  
XX PR 03-NOV-1999; 99US-0163280.  
XX PA (HUMA-) HUMAN GENOME SCI INC.  
XX PI Ruben SM, Barash SC, Birse CE, Rosen CA;  
XX WPI; 2001-235357/24.  
XX DR N-PSDB; AAH36465.  
XX PT Nucleic acids encoding 4277 human colon cancer-associated polypeptides,  
XX useful for preventing, diagnosing and/or treating colorectal cancers -  
XX Claim 11; Page 9172; 9803pp; English.  
XX AAH32943 to AAH37195 and AAH37198 represent human colon  
XX cancer-associated nucleic acid molecules (N) and proteins (P), where  
XX the proteins are collectively known as colon cancer antigens. The colon  
XX cancer antigens have cytostatic activity and can be used in gene  
XX therapy and vaccine production. N and P may be used in the prevention,  
XX diagnosis and treatment of diseases associated with inappropriate P  
XX expression. For example, N and P may be used to treat disorders  
XX associated with decreased expression by rectifying mutations or deletions  
XX in a patient's genome that affect the activity of P by expressing  
XX inactive proteins or to supplement the patients own production of P.  
XX Additionally, N may be used to produce the colon cancer-associated Ps,  
XX by inserting the nucleic acids into a host cell and culturing the cell  
XX to express the proteins. N and P can be used in the prevention, diagnosis  
XX and treatment of colorectal carcinomas and cancers. AAH37196 to AAH37204  
XX and AAH37789 represent sequences used in the exemplification of the  
XX present invention.  
XX N.B. Pages 666 to 682 and page 7053 of the sequence listing were  
XX missing at time of publication, meaning no sequences are present for  
XX SEQ ID NO:1027 to 1052, 7921 and 7922.  
XX Sequence 32 AA;  
XX Query Match 77.4%; Score 24; DB 22; Length 32;  
XX Best Local Similarity 83.3%; Pred. No. 91;  
XX Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;  
XX QY 1 ALNVLM 6  
XX :|||||  
XX 22 SLNVLM 27  
XX RESULT 39  
XX ABP10982  
XX ID ABP10982 standard; Protein; 61 AA.  
XX AC ABP10982;  
XX XX 24-JUN-2002 (first entry)  
XX DT Human ORFX protein sequence SEQ ID NO:21946.  
XX DE  
XX Human; open reading frame; ORFX; gene therapy; cancer; cirrhosis;  
XX hyperproliferative disorder; psoriasis; benign tumour; haemorrhage;  
XX degenerative disorder; osteoarthritis; neurodegenerative disorder;  
XX cardiovascular disease; diabetes mellitus; systemic lupus erythematosus;  
XX hypertension; hypothyroidism; cholesterol ester storage disease;  
XX immune deficiency; immune disorder; infectious disease;  
XX autoimmune disorder; rheumatoid arthritis; autoimmune thyroiditis;  
XX myasthenia gravis.  
XX OS Homo sapiens.  
XX WO200192523-A2.  
XX PD 06-DEC-2001.  
XX PF 29-MAY-2001; 2001WO-US10836.  
XX PR 30-MAY-2000; 2000US-206132P.  
XX PR 29-AUG-2000; 2000US-228716P.  
XX PA (CURA-) CURAGEN CORP.  
XX PI Shimkets RA, Leach MD;  
XX WPI; 2002-106308/14.  
XX DR N-PSDB; ABN26734.  
XX PT Novel human polypeptides and polynucleotides useful for diagnosing,  
XX preventing and treating cardiovascular disease, neurodegenerative,  
XX hyperproliferative disorders and autoimmune disorders -  
XX Disclosure; SEQ ID 21946; 1037pp; English.  
XX The present invention describes substantially purified human proteins  
XX (referred to as open reading frame, ORFX, where X is 1-11491 (see Table 1  
XX in the specification). ABN15762 to ABN27252 encode the human ORFX for  
XX proteins given in ABP00010 to ABP11500. ORFX proteins are useful for  
XX treating or preventing a pathology associated with an ORFX-associated  
XX disorder in humans, and in the manufacture of a medicament for treating a  
XX syndrome associated with ORFX-associated disorder. ORFX polynucleotide  
XX sequences can be used in gene therapy. ORFX sequences can be used in the  
XX treatment of cancer, hyperproliferative disorders, disorders related to liver,  
XX psoriasis, benign tumours, keloid, degenerative disorders, cirrhosis of liver,  
XX osteoarthritis, neurodegenerative disorders, diabetes mellitus, systemic  
XX lupus erythematosus, hypertension, hypothyroidism, cholesterol ester  
XX storage disease, various immune deficiencies and disorders, rheumatoid  
XX diseases, autoimmune disorders such as multiple sclerosis, graft-versus-host  
XX arthritis, autoimmune thyroiditis, myasthenia gravis, rheumatoid  
XX disease and autoimmune inflammatory eye disease. ORFX proteins are also  
XX useful for treating burns, incisions, ulcers, for treating osteoporosis,  
XX bone degenerative disorders, or periodontal disease, and for gut  
XX protection or regeneration and treatment of lung or liver fibrosis,  
XX reperfusion injury in various tissues and conditions resulting from  
XX systemic cytokine damage.  
XX N.B. The sequence data for this patent did not form part of the printed  
XX specification, but was obtained in electronic format directly from WIPO  
XX at ftp.wipo.int/pub/published\_pct\_sequences.  
XX Query Match 77.4%; Score 24; DB 23; Length 61;  
XX Best Local Similarity 71.4%; Pred. No. 1.9e+02;  
XX Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;  
XX QY 1 ALNVLM 7  
XX :|||:  
XX 7 ALNVLA 13  
XX Db  
XX RESULT 40  
XX AAG91228  
XX ID AAG91228 standard; Protein; 64 AA.  
XX AC AAG91228;  
XX XX 26-SEP-2001 (first entry)  
XX DT C glutamicum protein fragment SEQ ID NO: 4982.  
XX DE  
XX Coryneform bacterium; amino acid synthesis; vitamin; saccharide;  
XX organic acid synthesis.  
XX KW

```

XX OS Corynebacterium glutamicum.
XX PN EP1108790-A2.
XX PD 20-JUN-2001.
XX PF 18-DEC-2000; 2000EP-0127688.
XX PR 16-DEC-1999; 99JP-0377484.
XX PR 07-APR-2000; 2000JP-0159162.
XX PR 03-AUG-2000; 2000JP-0280988.
XX PA (KYOW ) KYOWA HAKKO KOGYO KK.
XX PI Nakagawa S, Mizoguchi H, Ando S, Hayashi M, Ochiai K, Yokoi H;
PI Tateishi N, Senoh A, Ikeda M, Ozaki A;
XX DR WPI; 2001-376931/40.
XX DR N-PSDB; AAH66447.
XX PT Novel polynucleotides derived from Coryneform bacteria, for identifying
PT mutation point of a gene, measuring expression of a gene, analysing
PT expression profile or pattern of a gene and identifying homologous gene
PT
XX PS Claim 17; SEQ ID NO: 4982; 246pp + Sequence Listing; English.
XX CC The present invention provides a number of nucleotide and protein
CC sequences from the Coryneform bacterium Corynebacterium glutamicum. These
CC are useful for identifying the mutation point of a gene derived from a
CC mutant of coryneform bacterium, measuring expression amount and
CC analysing the expression profile or expression pattern of a gene derived
CC from Coryneform bacterium, and identifying a homologue of a gene derived
CC from coryneform bacterium. Coryneform bacteria are useful for producing
CC amino acids, nucleic acids, vitamins, saccharides and organic acids,
CC particularly L-lysine. The present sequence is a protein described
CC in the exemplification of the invention.
CC Note: The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from the
CC European Patent Office.
XX SQ Sequence 64 AA;

Query Match 77.4%; Score 24; DB 22; Length 64;
Best Local Similarity 83.3%; Pred. No. 2e+02;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Oy 1 ALNVLM 6
Db 24 ALNVLL 29
|||||

RESULT 41
AAU17920
ID AAU17920 standard; Protein; 65 AA.
XX AC AAU17920;
XX DT 07-NOV-2001 (first entry)
XX DE
XX DE Novel human respiratory antigen #236.
XX KW Human; respiratory antigen; respiratory disorder; throat disorder;
XX KW lung disorder; nose disorder; lung cancer; gene therapy; cytostatic;
XX KW anti allergic; anti asthmatic; anti inflammatory; olfactory;
XX KW respiratory active.
XX OS Homo sapiens.
XX PN WO200155448-A1.
XX PD 02-AUG-2001.

```

```

XX PF 17-JAN-2001; 2001WO-US01333.
XX PR 31-JAN-2000; 2000US-0179065.
XX PR 04-FEB-2000; 2000US-0180628.
XX PR 24-FEB-2000; 2000US-0184664.
XX PR 02-MAR-2000; 2000US-0186350.
XX PR 16-MAR-2000; 2000US-0189874.
XX PR 17-MAR-2000; 2000US-0190076.
XX PR 18-APR-2000; 2000US-0198123.
XX PR 19-MAY-2000; 2000US-0205515.
XX PR 07-JUN-2000; 2000US-0209467.
XX PR 28-JUN-2000; 2000US-0214886.
XX PR 30-JUN-2000; 2000US-0215135.
XX PR 07-JUL-2000; 2000US-0216647.
XX PR 11-JUL-2000; 2000US-0217487.
XX PR 14-JUL-2000; 2000US-0218290.
XX PR 26-JUL-2000; 2000US-0220963.
XX PR 26-JUL-2000; 2000US-0220964.
XX PR 14-AUG-2000; 2000US-0224518.
XX PR 14-AUG-2000; 2000US-0224519.
XX PR 14-AUG-2000; 2000US-0225213.
XX PR 14-AUG-2000; 2000US-0225214.
XX PR 14-AUG-2000; 2000US-0225266.
XX PR 14-AUG-2000; 2000US-0225267.
XX PR 14-AUG-2000; 2000US-0225268.
XX PR 14-AUG-2000; 2000US-0225270.
XX PR 14-AUG-2000; 2000US-0225447.
XX PR 14-AUG-2000; 2000US-0225757.
XX PR 14-AUG-2000; 2000US-0225758.
XX PR 18-AUG-2000; 2000US-0225759.
XX PR 18-AUG-2000; 2000US-0226279.
XX PR 22-AUG-2000; 2000US-0226681.
XX PR 22-AUG-2000; 2000US-0226868.
XX PR 22-AUG-2000; 2000US-0227182.
XX PR 23-AUG-2000; 2000US-0227009.
XX PR 30-AUG-2000; 2000US-0228924.
XX PR 01-SEP-2000; 2000US-0229287.
XX PR 01-SEP-2000; 2000US-0229343.
XX PR 01-SEP-2000; 2000US-0229344.
XX PR 01-SEP-2000; 2000US-0229345.
XX PR 05-SEP-2000; 2000US-0229509.
XX PR 05-SEP-2000; 2000US-0229513.
XX PR 06-SEP-2000; 2000US-0230437.
XX PR 06-SEP-2000; 2000US-0230438.
XX PR 08-SEP-2000; 2000US-0231242.
XX PR 08-SEP-2000; 2000US-0231243.
XX PR 08-SEP-2000; 2000US-0231244.
XX PR 08-SEP-2000; 2000US-0231413.
XX PR 08-SEP-2000; 2000US-0231414.
XX PR 08-SEP-2000; 2000US-0231414.
XX PR 08-SEP-2000; 2000US-0232080.
XX PR 08-SEP-2000; 2000US-0232081.
XX PR 12-SEP-2000; 2000US-0231968.
XX PR 14-SEP-2000; 2000US-0232397.
XX PR 14-SEP-2000; 2000US-0232398.
XX PR 14-SEP-2000; 2000US-0232399.
XX PR 14-SEP-2000; 2000US-0232400.
XX PR 14-SEP-2000; 2000US-0232401.
XX PR 14-SEP-2000; 2000US-0233063.
XX PR 14-SEP-2000; 2000US-0233064.
XX PR 14-SEP-2000; 2000US-0233065.
XX PR 21-SEP-2000; 2000US-0234223.
XX PR 21-SEP-2000; 2000US-0234274.
XX PR 23-SEP-2000; 2000US-0234997.
XX PR 25-SEP-2000; 2000US-0234998.
XX PR 26-SEP-2000; 2000US-0235484.
XX PR 27-SEP-2000; 2000US-0235834.
XX PR 27-SEP-2000; 2000US-0235836.
XX PR 29-SEP-2000; 2000US-0236327.
XX PR 29-SEP-2000; 2000US-0236367.
XX PR 29-SEP-2000; 2000US-0236368.

```

XX Isolated polypeptide for treating, preventing and/or prognosing  
PT disorders related to the respiratory system including respiratory  
PT cancers and also for testing and detection e.g. diagnosis -  
XX  
PS Claim 11; SED ID No 538; 546pp; English.  
XX The present invention relates to the isolation of novel human  
CC respiratory antigens, and cDNA (AA827869-AAS28159) and genomic  
CC sequences encoding for these polypeptides. The sequences of the  
CC invention are useful for preventing, treating and/or prognosing  
CC disorders related to the respiratory system including throat  
CC disorders (e.g. vocal cord paralysis, tonsillitis, and laryngitis),  
CC lung disorders e.g. pneumonia, allergic disorders e.g. asthma,  
CC pleurisy, cystic fibrosis, emphysema, nose disorders and cancers of  
CC the respiratory tissues e.g. lung cancer. The polynucleotide sequences  
CC of the invention are useful in gene therapy and antisense therapy.  
CC AAU17685-AAU17975 represent novel human respiratory antigens.  
CC Note: The sequence data for this patent did not form part of the printed  
CC specification, but was obtained in electronic format directly from WIPO  
CC at ftp.wipo.int/pub/published\_pct\_sequences.  
XX  
SQ Sequence 65 AA;  
Query Match 77.4%; Score 24; DB 22; Length 65;  
Best Local Similarity 66.7%; Pred. No. 2.1e+02;  
Matches 4; Conservative 2; Mismatches 0; Indels 0; Gaps 0;  
QY 2 LNVLMNA 7  
Db 28 LNIILMA 33  
RESULT 42  
AAS56878  
ID AAS56878 standard; Protein; 71 AA.  
XX AAS56878;  
XX 13-MAR-2001 (first entry)  
XX Human prostate cancer antigen protein sequence SEQ ID NO:1456.  
XX Human; prostate cancer; prostate cancer antigen; detection; diagnosis;  
XX neuroprotective; cytostatic; cardioactive; immunomodulatory; muscular;  
XX vulnary; gastrointestinal; nephrotropic; antiinfective; gynaecological;  
XX antibacterial; gene therapy; neural; immune; reproductive; renal;  
XX gastrointestinal; pulmonary; cardiovascular; proliferative disorder;  
XX wound; infectious disease.  
XX Homo sapiens.  
XX WO2000055174-A1.  
XX 21-SEP-2000.  
XX 08-MAR-2000; 2000WO-US05988.  
XX 12-MAR-1999; 99US-0124270.  
XX (HUMA-) HUMAN GENOME SCI INC.  
XX (ROSE/) ROSEN C A.  
XX Rosen CA, Ruben SM;  
XX WPI; 2000-587513/55.  
XX N-PSDB; AAF16081.  
XX Prostate cancer associated gene sequences, referred to as prostate  
XX cancer antigens, useful for treatment, prevention, and diagnosis of  
XX disorders such as prostate cancer -  
XX Claim 11; Page 1890; 2338pp; English.  
PS

29-SEP-2000; 2000US-0236369.  
29-SEP-2000; 2000US-0236370.  
02-OCT-2000; 2000US-0236802.  
02-OCT-2000; 2000US-0237037.  
02-OCT-2000; 2000US-0237038.  
02-OCT-2000; 2000US-0237039.  
02-OCT-2000; 2000US-0237040.  
13-OCT-2000; 2000US-0239335.  
13-OCT-2000; 2000US-0239937.  
20-OCT-2000; 2000US-0240960.  
20-OCT-2000; 2000US-0241221.  
20-OCT-2000; 2000US-0241785.  
20-OCT-2000; 2000US-0241786.  
20-OCT-2000; 2000US-0241787.  
20-OCT-2000; 2000US-0241808.  
20-OCT-2000; 2000US-0241809.  
20-OCT-2000; 2000US-0241826.  
01-NOV-2000; 2000US-0244617.  
08-NOV-2000; 2000US-0246474.  
08-NOV-2000; 2000US-0246475.  
08-NOV-2000; 2000US-0246476.  
08-NOV-2000; 2000US-0246477.  
08-NOV-2000; 2000US-0246478.  
08-NOV-2000; 2000US-0246523.  
08-NOV-2000; 2000US-0246524.  
08-NOV-2000; 2000US-0246525.  
08-NOV-2000; 2000US-0246526.  
08-NOV-2000; 2000US-0246527.  
08-NOV-2000; 2000US-0246528.  
08-NOV-2000; 2000US-0246532.  
08-NOV-2000; 2000US-0246609.  
08-NOV-2000; 2000US-0246610.  
08-NOV-2000; 2000US-0246611.  
08-NOV-2000; 2000US-0246613.  
17-NOV-2000; 2000US-0249207.  
17-NOV-2000; 2000US-0249208.  
17-NOV-2000; 2000US-0249209.  
17-NOV-2000; 2000US-0249210.  
17-NOV-2000; 2000US-0249211.  
17-NOV-2000; 2000US-0249212.  
17-NOV-2000; 2000US-0249213.  
17-NOV-2000; 2000US-0249214.  
17-NOV-2000; 2000US-0249215.  
17-NOV-2000; 2000US-0249216.  
17-NOV-2000; 2000US-0249217.  
17-NOV-2000; 2000US-0249218.  
17-NOV-2000; 2000US-0249244.  
17-NOV-2000; 2000US-0249245.  
17-NOV-2000; 2000US-0249264.  
17-NOV-2000; 2000US-0249265.  
17-NOV-2000; 2000US-0249297.  
17-NOV-2000; 2000US-0249299.  
17-NOV-2000; 2000US-0249300.  
01-DEC-2000; 2000US-0250160.  
01-DEC-2000; 2000US-0250391.  
05-DEC-2000; 2000US-0251030.  
05-DEC-2000; 2000US-0251388.  
05-DEC-2000; 2000US-0256719.  
06-DEC-2000; 2000US-0251479.  
08-DEC-2000; 2000US-0251856.  
08-DEC-2000; 2000US-0251868.  
08-DEC-2000; 2000US-0251869.  
08-DEC-2000; 2000US-0251989.  
08-DEC-2000; 2000US-0251990.  
11-DEC-2000; 2000US-0254097.  
05-JAN-2001; 2001US-0259678.  
(HUMA-) HUMAN GENOME SCI INC.  
Rosen CA, Barash SC, Ruben SM;  
WPI; 2001-476224/51.  
N-PSDB; AAS28104.

XX AAF15566 to AAF16505 encode the human prostate cancer associated  
 CC proteins, called prostate cancer antigens, given in AAB56363 to AAB57302.  
 CC The prostate cancer antigens can have neuroprotective, cytostatic,  
 CC cardioactive, immunomodulatory, muscular, vulnary, gastrointestinal,  
 CC nephrotropic, antiinfective, gynaecological and antibacterial activities,  
 CC and can be used in gene therapy. The prostate cancer antigen  
 CC polynucleotides may be used for detection of prostate cancer, chromosome  
 CC identification, as chromosome markers, and for numerous other diagnostic  
 CC or research purposes. The prostate cancer antigens may be used to treat  
 CC disorders such as neural, immune, muscular, reproductive,  
 CC gastrointestinal, pulmonary, cardiovascular, renal, and proliferative  
 CC disorders, wounds, and infectious diseases. AAF16506 to AAF16514 to  
 CC AAB57303 represent sequences used in the exemplification of the present  
 CC invention.

XX SQ Sequence 71 AA;  
 Query Match 77.4%; Score 24; DB 21; Length 71;  
 Best Local Similarity 66.7%; Pred. No. 2.3e+02;  
 Matches 4; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 2 LNLVMA 7  
 Db 35 LNLVMA 40  
 ||:|  
 ||:|

RESULT 43  
 AAU27363  
 ID AAU27363 standard; Protein; 75 AA.  
 XX AC AAU27363;  
 XX -DT 18-DEC-2001 (first entry)  
 XX DE Novel bone marrow polypeptide #41.  
 XX KW Human; bone marrow; inflammation; arthritis; nephritis; Crohn's disease;  
 KW ischaemia-reperfusion injury; shock; sepsis; haematopoiesis; bone growth;  
 KW cancer; metastasis; transgenic animal; nerve regeneration; neuropathy;  
 KW Alzheimer's disease; Parkinson's disease; Huntington's disease; ALS;  
 KW amyotrophic lateral sclerosis; lymphoid cell disorder; platelet disorder;  
 KW thrombocytopenia; burn; ulcer; osteoporosis; periodontal disease; SCID;  
 KW lung fibrosis; liver fibrosis; immune deficiency; autoimmune disorder;  
 KW severe combined immunodeficiency; infection; multiple sclerosis.

XX OS Homo sapiens.  
 XX PN WO200164840-A2.  
 XX PD 07-SEP-2001.  
 XX PF 28-FEB-2001; 2001WO-US06509.  
 XX PR 28-FEB-2000; 2000US-0515126.  
 XX PR 18-MAY-2000; 2000US-0577409.  
 XX PR 30-NOV-2000; 2000US-0250583.  
 XX PA (HYSE-) HYSEQ INC.  
 XX PI Ford JE, Boyle BJ, Tang YT, Liu C, Drmanac RT;  
 XX WPI; 2001-582153/65.  
 XX DR N-PSDB; AAS43306.  
 XX Novel bone marrow-expressed nucleic acids and polypeptides for  
 PT research, diagnosis and treatment of haematopoietic, autoimmune,  
 PT inflammatory disorders and cancer and for use in stem cell survival,  
 PT bone marrow and remodeling -  
 XX Claim 10; Page 170; 228pp; English.  
 PS The invention relates to novel isolated bone marrow-expressed  
 XX

CC polynucleotide (I) (or its complement) comprising a sequence selected  
 CC from 150 sequences of defined base pair sequences given in the  
 CC specification, its mature coding portion or active domain coding portion.  
 CC (I) provided as a collection on a nucleic acid array is useful  
 CC for detecting full-matches or mismatches to any one of the  
 CC polynucleotides in the collection. (I) and its encoded polypeptides (II)  
 CC are useful for treating inflammatory conditions such as arthritis,  
 CC nephritis, Crohn's disease, ischaemia-reperfusion injury, shock, sepsis,  
 CC immune responses, and are involved in increasing haematopoiesis, stem  
 CC cell survival, bone growth and remodeling. (II) is involved in cancer  
 CC cell generation and proliferation in metastasis. (I), (II), and modulators  
 CC of (I) are useful for prophylaxis or treatment of cancer. (I) is also  
 CC useful for creating transgenic animals useful for studying the in vivo  
 CC activities of the polypeptide as well as for studying modulators of the  
 CC polypeptides. (II) induces the proliferation of neural cells and  
 CC regeneration of nerve and brain tissue and is useful for the  
 CC treatment of central and peripheral nervous system diseases and  
 CC neuropathies, such as Alzheimer's, Parkinson's disease, Huntington's  
 CC disease and amyotrophic lateral sclerosis. (II) is involved in  
 CC chemotactic or chemokinetic activity, regulation of haematopoiesis and  
 CC is useful for treating myeloid or lymphoid cell disorders, platelet  
 CC disorders such as thrombocytopenia and for regeneration of bone,  
 CC cartilage, tendon, ligament and/or nerve tissue growth, and in tissue  
 CC repair, healing of burns, incisions, ulcers, for treating osteoporosis,  
 CC osteoarthritis, bone degenerative disorders or periodontal disease.  
 CC (II) is also useful for gut protection or regeneration and treatment of  
 CC lung or liver fibrosis, reperfusion injury in various tissues, various  
 CC immune deficiencies and disorders including severe combined  
 CC immunodeficiency (SCID), bacterial or fungal infections, and autoimmune  
 CC disorders e.g. multiple sclerosis. AAU27323-AAU27472 represent human bone  
 CC marrow polypeptide sequences of the invention.

XX SQ Sequence 75 AA;  
 Query Match 77.4%; Score 24; DB 22; Length 75;  
 Best Local Similarity 57.1%; Pred. No. 2.5e+02;  
 Matches 4; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 1 ALNVLMA 7  
 Db 41 SVNILMA 47  
 :|:|  
 :|:|

RESULT 44  
 AAEO9109  
 ID AAEO9109 standard; peptide; 7 AA.  
 XX AC AAEO9109;  
 XX -DT 15-NOV-2001 (first entry)  
 XX DE Bacteriophage Type III peptide #1, used in the invention.  
 XX KW Vaccine; therapy; autoimmune disorder; Epstein-Barr virus; EBV; anaemia;  
 KW polymyositis; systemic lupus erythematosus; SLE; rheumatoid arthritis;  
 KW Sjogren's syndrome; diabetes mellitus; adenitis; multiple sclerosis;  
 KW demyelinating disease; Hashimoto's thyroiditis; autoimmune infertility;  
 KW hypoparathyroidism; primary biliary cirrhosis; ankylosing spondylitis;  
 KW inflammatory bowel disease; Addison's disease; thyroiditis; filariasis;  
 KW Graves' disease; ulcerative colitis; dermatomyositis; myasthenia gravis;  
 KW Crohn's disease; CREST syndrome; autoimmune cystitis; glomerulonephritis;  
 KW polyarthritis nodosa; hepatitis; atopic rhinitis; Goodpasture's syndrome;  
 KW sarcoidosis; rheumatic fever; anti-phospholipid syndrome; farmer's lung;  
 KW Cushing's syndrome; bird-fancier's lung; alveolitis; erythema nodosum;  
 KW pyoderma gangrenosum; fibromyalgia; Kawasaki's disease; cardiomyopathy;  
 KW Sampter's syndrome; asthma; polymyalgia rheumatica; psoriasis; arthritis;  
 KW erythroblastosis foetalis; cyclitis; IGA nephropathy; Hodgkin's lymphoma;  
 KW renal cell carcinoma; eosinophilia; immunosuppressive; ophthalmological;  
 KW thymimetic; neuroprotective; cytostatic; nephrotropic; anti-allergic;  
 KW dengue; antitumor; vasotropic; antipyretic; hepatotropic.  
 XX OS Bacteriophage.



GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: February 17, 2004, 10:50:12 ; Search time 42.4752 Seconds  
(without alignments)  
182.261 Million cell updates/sec

**Title:** US-09-900-147-6

Perfect score:

Sequence: 1 YDALNVLAMNIIISKEKEIKWIGLPTNSA 30

Scoring table: BLOSUM62

Gapop 10.0 ; Gapext 0.5

Searched: 830525 seqs. 258052604 residues

Total number of hits satisfying chosen parameters: 146963

Minimum DB seq length: 0

Maximum DB seq length:	0
Maximum DB seq length:	100

Post-processing: Minimum Match 0%

Fast-processing: Minimum Match 0%  
Maximum Match 100%

Maximum Match 100%  
Listing first 45 summaries

Database :

```

SPREMBL_23.*
1: sp archaea.*
2: sp bacteria.*
3: sp fungi.*
4: sp human.*
5: sp invertebrate.*
6: sp mammal.*
7: sp mhc.*
8: sp organelle.*
9: sp phase.*
10: sp plant.*
11: sp rodent.*
12: sp virus.*
13: sp vertebrae.*
14: sp unclassified.*
15: sp rivirus.*
16: sp bacteriap.*
17: sp archaeap.*

```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query %			DB	ID	Description
		Match	Length				
1	45.5	29.9	84	2	Q50148	mycobacteri	
2	44	28.9	89	10	Q23974	arabidopsis	
3	43	28.3	75	10	Q8LR77	oryza sativ	
4	43	28.3	76	16	Q98R59	mycoplasma	
5	42	27.6	61	11	Q61910	mus muscula	
6	42	27.6	61	12	Q57258	vaccinia vi	
7	42	27.6	76	17	Q97X12	sulfolobus	
8	42	27.6	76	12	Q9JFT1	ectromelia	
9	42	27.6	96	16	Q99TQ4	staphylococ	
10	42	27.3	97	16	Q8R895	thermoanaer	
11	41.5	27.3	74	17	Q8PWU3	methanobac	
12	41	27.0	75	16	Q8E9Y5	shewanella	
13	41	27.0	100	12	Q9DHG3	yaba-like d	
14	40.5	26.6	65	2	Q9R1G5	yersinia en	
15	40	26.3	69	10	Q9S895	fagopyrium e	
16	40	26.3	69	10	Q9S896	fagopyrium e	

## ALIGNMENTS

## RESULT 1

Q50148	PRELIMINARY;	PRT;	84 AA.
ID	Q50148		
AC	Q50148;		
DT	01-NOV-1996	(TEMBLrel. 01, Created)	
DT	01-NOV-1996	(TEMBLrel. 01, Last sequence update)	
DT	01-NOV-1996	(TEMBLrel. 01, Last annotation update)	
DE	U2961.		
OS	Mycobacterium leprae.		
OC	Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;		
OC	Corynebacterinae; Mycobacteriaceae; Mycobacterium.		
OX	NCBI_TaxID=1769;		
	[1]		
RN	SEQUENCE FROM N.A.		
RP			

Query Match 29.9%: Score 45.5: DB 2: Length 84:

Qy 8 MAMNIIISKEKKEIKWIGLPTNS 29  
||| ::| : | : ||| ::  
Db 1 MAM-LMSRIESEASWMGLPVDA 21

## RESULT 2

023974	PRELIMINARY;	PRT;	89 AA.
AD	O23974		
ID	AC	O23974;	
DT	DT	01-JAN-1998	(TrenBurel. 05, Created)
DT	DT	01-JAN-1998	(TrenBurel. 05, Last sequence update)
DT	DT	01-OCT-2002	(TrenBurel. 22, Last annotation update)

```

Q98R59;
AC 01-OCT-2001 (TREMBLrel. 18, Created)
DT 01-OCT-2001 (TREMBLrel. 18, Last sequence update)
DT 01-OCT-2001 (TREMBLrel. 18, Last annotation update)
DE Transposase for insertion sequence element IS1138.
DE MYPU 1510.
GN MYCOPlasma pulmonis.
OS Bacteria; Firmicutes; Mollicutes; Mycoplasmataceae; Mycoplasma
OC NCBI_TaxID=2107;
OX [1]
RN SEQUENCE FROM N.A.
RP STRAIN=UAB CTIP:
PC MEDLINE=21267165; PubMed=11353084;
RX Chambaud I., Heilig R., Ferris S., Barbe V., Samson D., Galisson
RA Moszer I., Dybvik K., Wroblewski H., Viari A., Rocha E.P.C.,
RA Blanchard A.;
RA "The complete genome sequence of the murine respiratory pathogen
RT Mycoplasma pulmonis."
RL Nucleic Acids Res. 29:2145-2153(2001).
DR EMBL; AL445563; CAC1324.1; -.
DR Mypulist; MYPU 1510; -.
KW Complete proteome.
SQ SEQUENCE 78 AA; 9386 MW; 05018A7BF7D87E46 CRC64;
Query Match 28.3%; Score 43; DB 16; Length 78;
Best Local Similarity 45.0%; Pred. No. 94;
Matches 9; Conservative 5; Mismatches 6; Indels 0;

```

Query Match 28.3%; Score 43; DB 16; Length 78;  
Best Local Similarity 45.0%; Pred. No. 94;  
Matches 9; Conservative 5; Mismatches 6; Indels 0; Gaps 0;

Qy 4 LNVLMAMNIIISKEKEIKWI 23  
||| : : ||| : : |||

p6 9 LNVKLILDHISKNKFDKEWI 28

RESULT 5						
ID	Q61910	PRELIMINARY;	PRT;	61 AA.		
AC	Q61910;					
DT	01-NOV-1996	(TREMBLrel. 01, Created)				
DT	01-NOV-1996	(TREMBLrel. 01, Last sequence update)				
DT	01-OCT-2002	(TREMBLrel. 22, Last annotation update)				
DE	Mammary transforming protein.					
GN	PTA15 OR MAT1.					
OS	Mus musculus (Mouse).					
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;					
OC	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.					
OX	NCBI_TaxID=10090;					
RN	[1]					
RP	SEQUENCE FROM N.A.					
RC	STRAIN=BALB/c; TISSUE=Breast tumor;					
RX	MEDLINE=95024046; PubMed=7937892;					
RA	Bera T.K., Guzman R.C., Miyamoto S., Panda D.K., Sasaki M., Hanyu K.,					
RA	Enami J., Nandi S.;					
RT	"Identification of a mammary transforming gene (MAT1) associated with					
RT	mouse mammary carcinogenesis.";					
EL	Proc. Natl. Acad. Sci. U.S.A. 91:9789-9793(1994).					
DR	EWBL; L31958; AAC37665.1; --					
DR	MGI; MGI:104799; Peals.					
SQ	SEQUENCE 61 AA; 7137 MW; DDC1F91842892171 CRC64;					
Query Match	27.6%; Score 42; DB 11; Length 61;					
Best Local Similarity	33.3%; Pred. NO. 1e+02; 2. Tadele					
						2. Gaps

Query Match	27.6%	Score 42;	DB 11;	Length 61;
Best Local Similarity	33.3%	Pred. No. 1e+02;		
Matches	7;	Conservative	9;	Mismatches
			3;	Indels
Y	6	VLMAMNIISKEKBEIKWIGLP	26	
		: : :   : :   :		
	14	VVFSTNVIISRPERE--WEGMP	32	

```

6 VLAMNIIISKEKEIKWIGLP 26
  | : : | : : : | : | : |
14 VVRSNIIJSRPERE--WEGMP 32

```

RESULT 6		
O57258		
ID O57258	PRELIMINARY;	PRT;
AC O57258;		
nm 01-JUN-1998	(T=EMBLrel. 06, Created)	74 AA.



```
DT 01-JUN-1998 (TrEMBLrel. 06, Last sequence update)
DE 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE Putative 8.5k protein.
GN MVA179R.
OS Vaccinia virus (strain Ankara).
OC Viruses; dsDNA viruses, no RNA stage; Poxviridae; Chordopoxvirinae;
OC Orthopoxvirus.
OX NCBI_TaxID=126794;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Ankara;
RA Antoine G., Scheiflinger F., Falkner F.G., Dorner F.;
RT "The complete genomic sequence of the Modified Vaccinia Ankara (MVA)
RL strain.";
RL Submitted (MAR-1997) to the EMBL/GenBank/DBJ databases.
DR EMBL; U94848; AAB96552.1; -.
SQ SEQUENCE 74 AA; 8501 MW; 7152A06380829716 CRC64;

Query Match 27.6%; Score 42; DB 12; Length 74;
Best Local Similarity 52.6%; Pred. No. 1.3e+02;
Matches 10; Conservative 2; Mismatches 7; Indels 0; Gaps 0;

QY 2 DALNVLMMNNIISKKEIK 20
Db | | | | | | | | | | | | | | | | | | | |
6 DVTNVEDIMNEIDREKEEI 24

RESULT 7
Q97XL2 PRELIMINARY; PRT; 76 AA.
AC Q97XL2;
DT 01-OCT-2001 (TrEMBLrel. 18, Created)
DT 01-OCT-2001 (TrEMBLrel. 18, Last sequence update)
DE 01-OCT-2001 (TrEMBLrel. 18, Last annotation update)
DE First ORF in transposon ISC1491.
GN SS08760.
OS Sulfolobus solfataricus.
OC Archaea; Crenarchaeota; Thermoprotei; Sulfolobales; Sulfolobaceae;
OC Sulfolobus.
OX NCBI_TaxID=2287;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC 35092 / DSM 1617 / P2;
RX MEDLINE=21332296; PubMed=11427726;
RA She Q., Singh R.K., Confalonieri F., Zivanovic Y., Allard G.,
RA Aweez M.J., Chan-Weiler C.C.-Y., Clausen I.G., Curtis B.A.,
RA De Moors A., Erauso G., Fletcher C., Gordon P.M.K.,
RA Heikamp-de Jong I., Jeffries A.C., Kozera C.J., Medina N., Peng X.,
RA Thi-Ngoc H.P., Redder P., Schenk M.E., Theriault C., Tolstrup N.,
RA Charlebois R.L., Doolittle W.F., Duguet M., Gaasterland T.,
RA Garrett R.A., Ragan M.A., Sengen C.W., Van der Oost J.;
RT "The complete genome of the crenarchaeon Sulfolobus solfataricus P2.";
RL Proc. Natl. Acad. Sci. U.S.A. 98:7835-7840(2001).
DR EMBL; AE006784; AAK41919.1; -.
KW Complete proteome.
SQ SEQUENCE 76 AA; 8966 MW; 67D288E4D56D431F CRC64;

Query Match 27.6%; Score 42; DB 17; Length 76;
Best Local Similarity 40.0%; Pred. No. 1.3e+02;
Matches 8; Conservative 7; Mismatches 5; Indels 0; Gaps 0;

QY 2 DALNVLMMNNIISKKEIK 21
Db | | | | | | | | | | | | | | | | | | | |
17 DKLTVVKAEIVITQEKVE 36

RESULT 8
Q9JFT1 PRELIMINARY; PRT; 87 AA.
AC Q9JFT1;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-OCT-2000 (TrEMBLrel. 15, Last annotation update)
```

```
DE CSR.
GN CSR.
OS Ectromelia virus.
OC Viruses; dsDNA viruses, no RNA stage; Poxviridae; Chordopoxvirinae;
OC Orthopoxvirus.
OX NCBI_TaxID=12643;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Moscow;
RA Chen N., Buller R.M.L., Wall E.M., Upton C.;
RT "Analysis of host response modifier ORFs of ectromelia virus, the
RT causative agent of mousepox.";
RL Virus Res. 66:155-173(2000).
DR EMBL; AF012825; AAC99564.1; -.
SQ SEQUENCE 87 AA; 9879 MW; ECAC2FA1023BAACE CRC64;

Query Match 27.6%; Score 42; DB 12; Length 87;
Best Local Similarity 52.6%; Pred. No. 1.5e+02;
Matches 10; Conservative 2; Mismatches 7; Indels 0; Gaps 0;

QY 2 DALNVLMMNNIISKKEIK 20
Db | | | | | | | | | | | | | | | | | | | |
19 DVTNVEDIMNEIDREKEEI 37

RESULT 9
Q99TQ4 PRELIMINARY; PRT; 96 AA.
AC Q99TQ4;
DT 01-JUN-2001 (TrEMBLrel. 17, Created)
DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DE 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE Hypothetical protein SAV1595 (Hypothetical protein MW1546).
GN SAV1595 OR SAL423 OR MW1546.
OS Staphylococcus aureus (strain Mu50 / ATCC 700699),
OS Staphylococcus aureus (strain N315), and
OS Staphylococcus aureus (strain MW2).
OC Bacteria; Firmicutes; Bacillales; Staphylococcus.
OX NCBI_TaxID=158878, 158879, 196620;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=MU50, and N315;
RX MEDLINE=21311952; PubMed=11418146;
RA Kuroda M., Ohta T., Uchiyama I., Baba T., Yuzawa H., Kobayashi I.,
RA Cui L., Oguchi A., Aoki K.-I., Nagai Y., Lian J.-Q., Ito T.,
RA Kanamori M., Matsumaru H., Maruyama A., Murakami H., Hoshoyama A.,
RA Mizutani-Uji Y., Takahashi N.K., Sawano T., Inoue R.-I., Kaito C.,
RA Sekimizu K., Hirakawa H., Kuhara S., Goto S., Yabuzaki J.,
RA Kanehisa M., Yamashita A., Oshima K., Furuya K., Yoshino C., Shiba T.,
RA Hattori M., Ogasawara N., Hayashi H., Hiramatsu K.;
RT "Whole genome sequencing of methicillin-resistant Staphylococcus
RT aureus.";
RL Lancet 357:1225-1240(2001).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=MW2;
RX MEDLINE=22040717; PubMed=12044378;
RA Baba T., Takeuchi F., Kuroda M., Yuzawa H., Aoki K.-I., Oguchi A.,
RA Nagai Y., Iwama N., Asano K., Naimi T., Kuroda H., Cui L.,
RA Yamamoto K., Hiramatsu K.;
RT "Genome and virulence determinants of high virulence community-
RT acquired MRSA.";
RL Lancet 359:1819-1827(2002).
DR EMBL; AP003362; BAB57757.1; -.
DR EMBL; AP003134; BAB42687.1; -.
DR EMBL; AP004827; BAB95411.1; -.
DR InterPro; IPR001890; UPF0044.
DR Pfam; PF01985; UPF0044; 1.
DR ProDom; PD010559; UPF0044; 1.
DR TIGRPFAMs; TIGR00253; TIGR00253; 1.
DR Hypothetical protein; Complete proteome.
KW Hypothetical protein; Complete proteome.
SQ SEQUENCE 96 AA; 11081 MW; E3C1845499AB3CA4 CRC64;
```

Query Match 27.3%; Score 41.5; DB 17; Length 74;  
Best Local Similarity 37.9%; Pred. No. 1.5e+02;  
Matches 11; Conservative 5; Mismatches 12; Indels 1; Gaps 1;

QY 2 DALNVLAMNIIISKEKEIKWIGLPTNSA 30  
DB 22 DRLLGLFALALKIMEYPTV-WIGLPSQTA 49

RESULT 12  
Q8EHV5 PRELIMINARY; PRT; 75 AA.  
AC Q8EHV5;  
DT 01-MAR-2003 (TRENBLrel. 23, Created)  
DT 01-MAR-2003 (TRENBLrel. 23, Last sequence update)  
DT 01-MAR-2003 (TRENBLrel. 23, Last annotation update)  
DE Conserved hypothetical protein.  
GN SO1069.  
OS Shewanella oneidensis.  
OC Bacteria; Proteobacteria; Gammaproteobacteria; Alteromonadales;  
OC Alteromonadales; Shewanella.  
OX NCBI\_TaxID=70863;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=MR-1;  
RX MEDLINE=22297686; PubMed=12368813;  
RA Heidelberg J.F., Paulsen I.T., Nelson K.E., Gaidos E.J., Nelson W.C.,  
RA Meyer T., Eisen J.A., Seshadri R., Ward N., Methe B., Clayton R.A.,  
RA DeBoy R.T., Dodson R.J., Durkin A.S., Haft D.H., Kolonay J.F.,  
RA Madupu R., Peterson J.D., Umayam L.A., White O., Wolf A.M.,  
RA Vamathevan J., Weidman J., Gill J., Utterback T.R., McDonald L.A.,  
RA Mueller J., Khouri H., Gill J., Venter J.C., Nealon K.H., Fraser C.M.,  
RA Feldlyum T.V., Smith H.O., Venter J.C., Nealon K.H., Fraser C.M.,  
RT "Genome sequence of the dissimilatory metal ion-reducing bacterium  
RT Shewanella oneidensis."  
RL Nat. Biotechnol. 20:1118-1123 (2002).  
DR EMBL; AE015551; AAN54141.1; -  
DR TIGR; SO1069; -  
KW Hypothetical protein; Complete proteome.  
SQ SEQUENCE 75 AA; 8592 MW; D243608FA8177F2C CRC64;

Query Match 27.0%; Score 41; DB 16; Length 75;  
Best Local Similarity 23.8%; Pred. No. 1.8e+02;  
Matches 5; Conservative 12; Mismatches 4; Indels 0; Gaps 0;

QY 5 NVLWAMNIIISKEKEIKWIGL 25  
DB 11 SVVVAISLWMSDIKLRWNL 31

RESULT 13  
Q9DHG3 PRELIMINARY; PRT; 100 AA.  
AC Q9DHG3;  
DT 01-MAR-2001 (TRENBLrel. 16, Created)  
DT 01-MAR-2001 (TRENBLrel. 16, Last sequence update)  
DT 01-DEC-2001 (TRENBLrel. 19, Last annotation update)  
DE 150R protein.  
GN 150R.  
OS Yaba-like disease virus (YLDV).  
OC Viruses; dsDNA viruses, no RNA stage; Poxviridae; Chordopoxvirinae;  
OC Yatapoxvirus.  
OX NCBI\_TaxID=132475;  
RN [1]  
RP SEQUENCE FROM N.A.  
RA Lee H.J.;  
RA Submitted (JUL-2000) to the EMBL/GenBank/DBJ databases.  
RN [2]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=21176366; PubMed=11277691;  
RA Lee H.J., Essani K., Smith G.L.;

Query Match 27.6%; Score 42; DB 16; Length 96;  
Best Local Similarity 40.0%; Pred. No. 1.6e+02;  
Matches 10; Conservative 6; Mismatches 9; Indels 0; Gaps 0;

QY 2 DALNVLAMNIIISKEKEIKWIGL 26  
DB 72 ELVQVIGSMIVYRESKENKEIPL 96

RESULT 10  
Q8R895 PRELIMINARY; PRT; 97 AA.  
AC Q8R895;  
DT 01-JUN-2002 (TRENBLrel. 21, Created)  
DT 01-JUN-2002 (TRENBLrel. 21, Last sequence update)  
DT 01-OCT-2002 (TRENBLrel. 22, Last annotation update)  
DE Hypothetical protein TTE2120.  
GN TTE2120.  
OS Thermoanaerobacter tengcongensis.  
OC Bacteria; Firmicutes; Clostridia; Thermoanaerobacteriales;  
OC Thermoanaerobacteriaceae; Thermoanaerobacter.  
OX NCBI\_TaxID=119072;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=MB4 / JCM 11007;  
RX MEDLINE=21992816; PubMed=11997336;  
RA Bao Q., Tian Y., Li W., Xu Z., Xuan Z., Hu S., Dong W., Yang J.,  
RA Chen Y., Xue Y., Xu Y., Lai X., Huang L., Dong X., Ma Y., Ling L.,  
RA Tan H., Chen K., Wang J., Yu J., Yang H.;  
RA "A complete sequence of T. tengcongensis genome."  
RL Genome Res. 12:689-700 (2002).  
DR EMBL; AE013159; AAM25286.1; -  
KW Hypothetical protein; Complete proteome.  
SQ SEQUENCE 97 AA; 11081 MW; 62B2381204D99127 CRC64;

Query Match 27.6%; Score 42; DB 16; Length 97;  
Best Local Similarity 40.7%; Pred. No. 1.6e+02;  
Matches 11; Conservative 4; Mismatches 8; Indels 4; Gaps 1;

QY 5 NVLWAMNIIISKEK---EIKWIGLPT 27  
DB 63 NILLTAYALISENVLIKIFWILLPT 89

RESULT 11  
Q8PWU3 PRELIMINARY; PRT; 74 AA.  
AC Q8PWU3;  
DT 01-OCT-2002 (TRENBLrel. 22, Created)  
DT 01-OCT-2002 (TRENBLrel. 22, Last sequence update)  
DT 01-OCT-2002 (TRENBLrel. 22, Last annotation update)  
DE Ferrous iron transport protein B.  
GN MW1483.  
OS Methanosarcina mazei (Methanosarcina frisia).  
OC Archaea; Euryarchaeota; Methanococci; Methanosarcinales;  
OC Methanosarcinaceae; Methanosarcina.  
OX NCBI\_TaxID=2209;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=Goel / Go1 / ATCC BAA-199 / DSM 3647 / OCM 88;  
RX MEDLINE=22120827; PubMed=12125824;  
RA Deppenmeier U., Johann A., Hartach T., Merkl R., Schmitz R.A.,  
RA Martinez-Arias R., Henne A., Wierze A., Baumer S., Jacobi C.,  
RA Brueggemann H., Lienard T., Christmann A., Boemcke M., Steckel S.,  
RA Bhattacharyya A., Lykidis A., Overbeek R., Klenk H.-P., Gunsalus R.P.,  
RA Fritz H.-J., Gottschalk G.;  
RA "The genome of Methanosarcina mazei: evidence for lateral gene  
RT transfer between Bacteria and Archaea."  
RL J. Mol. Microbiol. Biotechnol. 4:453-461 (2002).  
DR EMBL; AE013382; BAM31179.1; -  
KW Complete proteome.  
SQ SEQUENCE 74 AA; 8027 MW; E582EBF950B04F79 CRC64;

```

RT "The genome sequence of Yaba-like disease virus, a yatapoxvirus."
RL Virology 281:170-192(2001).
RN [3]
RA SEQUENCE FROM N.A.
RP Lee H.J.;
RL Thesis (2000), Sir William Dunn School of Pathology, University of.
DR ENBL; AJ293568; CAC21388.1; -.
SQ SEQUENCE 100 AA; 12130 MW; 5C5F098C03D6C337 CRC64;

Query Match 27.0%; Score 41; DB 12; Length 100;
Best Local Similarity 35.5%; Pred. No. 2.4e+02;
Matches 11; Conservative 6; Mismatches 8; Indels 6; Gaps 2;

QY 1 YDALNVLMANNIIS----KEKKE--IKWIGL 25
DB 59 YSFNFRLINKINPFYKEKEGILNLGM 89

RESULT 14
Q9RIGS PRELIMINARY; PRT; 65 AA.
AC Q9RIGS;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE Hypothetical 7.4 kDa protein (Fragment).
OS Yersinia enterocolitica.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Yersinia.
OX NCBI_TaxID=630;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=99426800; PubMed=10496882;
RA Bach S., Buchrieser C., Prentice M., Guiyoule A., Meadek T.,
RA Carniel E.;
RT "The high-pathogenicity island of Yersinia enterocolitica YE8081
RT undergoes low-frequency deletion but not precise excision, suggesting
RT recent stabilization in the genome."
RL Infect. Immun. 67:5091-5099(1999).
DR EMBL; AJ238284; CAB57394.1; -.
DR InterPro; IPR002178; PTS_EIIA_2.
DR Pfam; PF00359; PTS_EIIA_2; 1.
KW Hypothetical protein.
FT NON TER
FT NON TER
SQ SEQUENCE 65 AA; 7404 MW; 90E14C97C1216DF6 CRC64;

Query Match 26.6%; Score 40.5; DB 2; Length 65;
Best Local Similarity 42.9%; Pred. No. 1.9e+02;
Matches 9; Conservative 7; Mismatches 4; Indels 1; Gaps 1;

QY 1 YDALNVLMANNIISKEKEI 20
DB 10 WDRSNVNFIAIGVIAKEKEHI 30

RESULT 15
Q9S895 PRELIMINARY; PRT; 69 AA.
AC Q9S895;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
DE Bti-2-TRYPSIN inhibitor isoform.
OS Fagopyrum esculentum (Common buckwheat).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
OC Caryophyllidae; Caryophyllales; Polygonaceae; Fagopyrum.
OX NCBI_TaxID=3617;
RN [1]
RP SEQUENCE.
RX MEDLINE=97015347; PubMed=8862028;
RA Pandya M.J., Smith D.A., Farwood A., Gilroy J., Richardson M.;
RT "Complete amino acid sequences of two trypsin inhibitors from
RT buckwheat seed."
RL Phytochemistry 43:327-331(1996).
DR HSP; P19873; IHYM.
DR InterPro; IPR000864; Potato_inhibit.
DR Pfam; PF00280; potato_inhibit; 1.
DR ProDom; PD002604; Potato_inhibit; 1.
DR ProSite; PS00285; POTATO_INHIBITOR; 1.
SQ SEQUENCE 69 AA; 7748 MW; B2189CFC60C7E178 CRC64;

Query Match 26.3%; Score 40; DB 10; Length 69;
Best Local Similarity 40.9%; Pred. No. 2.4e+02;
Matches 9; Conservative 4; Mismatches 9; Indels 0; Gaps 0;

QY 9 AMNIIISKEKEIKWIGLPTNSA 30
DB 21 AAKIIENEDVRAIVLPEGSA 42

RESULT 16
Q9S896 PRELIMINARY; PRT; 69 AA.
AC Q9S896;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE Bti-1-7.634 kDa trypsin inhibitor isoform.
OS Fagopyrum esculentum (Common buckwheat).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
OC Caryophyllidae; Caryophyllales; Polygonaceae; Fagopyrum.
OX NCBI_TaxID=3617;
RN [1]
RP SEQUENCE.
RX MEDLINE=97015347; PubMed=8862028;
RA Pandya M.J., Smith D.A., Farwood A., Gilroy J., Richardson M.;
RT "Complete amino acid sequences of two trypsin inhibitors from
RT buckwheat seed."
RL Phytochemistry 43:327-331(1996).
DR HSP; P19873; IHYM.
DR InterPro; IPR000864; Potato_inhibit.
DR Pfam; PF00280; potato_inhibit; 1.
DR ProDom; PD002604; Potato_inhibit; 1.
DR ProSite; PS00285; POTATO_INHIBITOR; 1.
SQ SEQUENCE 69 AA; 7634 MW; ADC20CFC60D056D8 CRC64;

Query Match 26.3%; Score 40; DB 10; Length 69;
Best Local Similarity 40.9%; Pred. No. 2.4e+02;
Matches 9; Conservative 4; Mismatches 9; Indels 0; Gaps 0;

QY 9 AMNIIISKEKEIKWIGLPTNSA 30
DB 21 AAKIIENEDVRAIVLPEGSA 42

RESULT 17
Q9S9F3 PRELIMINARY; PRT; 69 AA.
AC Q9S9F3;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
DE Bti-1-PROTEASE inhibitor/trypsin inhibitor.
OS Fagopyrum esculentum (Common buckwheat).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
OC Caryophyllidae; Caryophyllales; Polygonaceae; Fagopyrum.
OX NCBI_TaxID=3617;
RN [1]
RP SEQUENCE.
RX MEDLINE=96013147; PubMed=7556606;
RA Belozersky M.A., Dunaevsky Y.E., Musolyamov A.X., Egorov T.A.;
RT "Complete amino acid sequence of the protease inhibitor from buckwheat

```

```
RC STRAIN=ATCC 35092 / DSM 1617 / P2;
RX MEDLINE=21332296; PubMed=11427726;
RA She Q., Singh R.K., Confalonieri F., Zivanovic Y., Allard G.,
RA HSP; P19873; IMIT.
RA InterPro; IPR000864; Potato_inhibit.
RA Pfam; PF00280; potato_inhibit; 1.
RA PRINTS; PR00292; POTATOINHTR.
RA ProDom; PD002604; Potato_inhibit; 1.
RA PROSITE; PS00285; POTATO_INHIBITOR; 1.
SQ SEQUENCE 69 AA; 7748 MW; A91887E76DCE178 CRC64;

Query Match 26.3%; Score 40; DB 10; Length 69;
Best Local Similarity 40.9%; Pred. No. 2.4e+02;
Matches 9; Conservative 4; Mismatches 9; Indels 0; Gaps 0;

QY 9 AMNIIISKEKEIKWIGLPTNSA 30.
Db 21 AAKIIENEDVRAIVLPESSA 42

RESULT 18
Q97YB4 PRELIMINARY; PRT; 76 AA.
AC Q97YB4;
DT 01-OCT-2001 (TrEMBLrel. 18, Created)
DT 01-OCT-2001 (TrEMBLrel. 18, Last sequence update)
DT 01-OCT-2001 (TrEMBLrel. 18, Last annotation update)
DE Partial ORF in transposon ISC1491.
DN SS07998.
OS Sulfolobus solfataricus.
OC Archaea; Crenarchaeota; Thermoprotei; Sulfolobales; Sulfolobaceae;
OC Sulfolobus.
OX NCBI_TaxID=2287;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC 35092 / DSM 1617 / P2;
RX MEDLINE=21332296; PubMed=11427726;
RA She Q., Singh R.K., Confalonieri F., Zivanovic Y., Allard G.,
RA Aweyez M.J., Chan-Weiher C.C.-Y., Clausen I.G., Curtis B.A.,
RA De Moors A., Erauso G., Fletcher C., Gordon P.M.K., Medina N., Peng X.,
RA Heikamp-de Jong I., Jeffries A.C., Kozera C.J., Theriault C., Tolstrup N.,
RA Thi-Ngoc H.P., Redder P., Schenk M.E., Theriault C., Tolstrup N.,
RA Charlebois R.L., Doolittle W.F., Duguet M., Gaasterland T.,
RA Garrett R.A., Ragan M.A., Sensen C.W., Van der Oost J.;
RT "The complete genome of the crenarchaeon Sulfolobus solfataricus P2."
RL Proc. Natl. Acad. Sci. U.S.A. 98:7835-7840(2001).
DR EMBL; AE006756; AAK41649.1; -.
KW Complete proteome.
SQ SEQUENCE 76 AA; 8950 MW; 8DEC9FE4D56D5A22 CRC64;

Query Match 26.3%; Score 40; DB 17; Length 76;
Best Local Similarity 40.0%; Pred. No. 2.6e+02;
Matches 8; Conservative 6; Mismatches 6; Indels 0; Gaps 0;

QY 2 DALNVLMMNIIISKEKEIK 21
Db 17 DKLTVAKAEIVTQEKREVE 36

RESULT 20
Q9JZ72 PRELIMINARY; PRT; 90 AA.
AC Q9JZ72;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-MAR-2002 (TrEMBLrel. 20, Last annotation update)
DE Hypothetical protein NMB1264.
DN NMB1264.
OS Neisseria meningitidis (serogroup B).
OC Bacteria; Proteobacteria; Betaproteobacteria; Neisseriales;
OC Neisseriaceae; Neisseria.
OX NCBI_TaxID=491;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=MC58 / Serogroup B;
RX MEDLINE=20175755; PubMed=10710307;
RA Tettelin H., Saunders N.J., Heidelberg J., Jeffries A.C., Nelson K.E.,
RA Eisen J.A., Ketchum K.A., Hood D.W., Peden J.F., Dodson R.J.,
RA Nelson W.C., Gwinn M.B., DeBoy R., Peterson J.D., Hickey E.K.,
RA Haft D.H., Salzberg S.L., White O., Fleischmann R.D., Dougherty B.A.,
RA Mason T., Ciecko A., Parksey D.S., Blair E., Citton H., Clark E.B.,
RA Cotton M.D., Utterback T.R., Khouri H., Qin H., Vamathevan J.,
RA Gill J., Scarlato V., Maignani V., Pizzi M., Grandi G., Sun L.,
RA Smith H.O., Fraser C.M., Moxon E.R., Rappuoli R., Venter J.C.;
RT "Complete genome sequence of Neisseria meningitidis serogroup B strain
MC58."
RL Science 287:1809-1815(2000).
DR EMBL; AE002475; AAF41641.1; -.
DR TIGR; NMB1264; -.
KW Hypothetical protein; Complete proteome.
SQ SEQUENCE 90 AA; 10394 MW; E8C7B7BD078225A CRC64;

Query Match 26.3%; Score 40; DB 16; Length 90;
Best Local Similarity 29.2%; Pred. No. 3.1e+02;
Matches 7; Conservative 8; Mismatches 9; Indels 0; Gaps 0;

QY 2 DALNVLMMNIIISKEKEIKWIGL 25
Db 53 NATHILKSYPEFAETKYLNWIGM 76

RESULT 21
Q95LH2 PRELIMINARY; PRT; 93 AA.
AC Q95LH2;
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-MAR-2002 (TrEMBLrel. 20, Last annotation update)
RP SEQUENCE FROM N.A.
```

DE	WAVE (Fragment).
GN	WAVE.
OS	Vibrio cholerae.
OC	Bacteria; Proteobacteria; Gammaproteobacteria; Vibrionales;
OC	Vibrionaceae; Vibrio.
OX	NCBI_TaxID=666;
RN	[1]
RP	SEQUENCE FROM N.A.
RC	STRAIN=V209;
RA	Nesper J., Kraiss A., Schild S., Blass J., Klose K.E., Bockemuehl J., Reidl J.;
RT	"Comparative and genetic analysis of the putative Vibrio cholerae LPS core oligosaccharide biosynthesis (wav) gene cluster.";
RL	Submitted (NOV-2001) to the EMBL/GenBank/DBJ databases.
DR	EMBL; AF443847; AAL77348.1; -.
FT	NON TER 1 1
SQ	SEQUENCE 74 AA; 8845 MW; 2BCD57E37F6FCE1 CRC64;
Query Match	26.0%; Score 39.5; DB 2; Length 74;
Best Local Similarity	38.2%; Pred. No. 3e+02;
Matches 13; Conservative	3; Mismatches 9; Indels 9; Gaps 2;
Qy	1 YDALNVLMANNIIISK---EKKEI-----KWIGL 25
Dd	29 YDLKDWLYNLGLISKNLTLDKKRIYRAYKKWTGL 62
RESULT 24	
Q8LLB0	PRELIMINARY; PRT; 72 AA.
ID	Q8LLB0
AC	Q8LLB0;
DT	01-OCT-2002 (TrEMBLrel. 22, Created)
DT	01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
DT	01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE	CIZE.
GS	CIZE.
ON	Hordeum vulgare (Barley).
OC	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC	Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Pooidae;
OC	Triticeae; Hordeum.
OX	NCBI_TaxID=4513;
RN	[1]
RP	SEQUENCE FROM N.A.
RC	STRAIN=cv. Morex;
RX	PubMed=12172030;
RA	Wei F., Wing R.A., Wise R.P.;
RT	"Genome dynamics and evolution of the Ma (powdery mildew) resistance locus in barley.";
RL	Plant Cell 14:1903-1917(2002).
DR	EMBL; AF427791; AAM22827.1; -.
DR	InterPro; IPR000864; Potato inhibit.
DR	Pfam; PF00280; potato inhibit; 1.
DR	PRINTS; PR00292; POTATOINHBT.
DR	ProDom; PD002604; Potato inhibit; 1.
SQ	SEQUENCE 72 AA; 7676 MW; 2030854E398438AA CRC64;
Query Match	25.7%; Score 39; DB 10; Length 72;
Best Local Similarity	37.0%; Pred. No. 3.5e+02;
Matches 10; Conservative	4; Mismatches 13; Indels 0; Gaps 0;
Qy	4 LNVLMMNNIIISKKKIKWIGLPTNSA 30
Dd	22 LTIKEAEIILKDKPEADIVVPVGSA 48
RESULT.25	
Q8EL48	PRELIMINARY; PRT; 73 AA.
ID	Q8EL48
AC	Q8EL48;
DT	01-MAR-2003 (TrEMBLrel. 23, Created)
DT	01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT	01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE	Hypothetical protein.

```

GN OB3383.
OS Oceanobacillus iheyensis.
OC Bacteria; Firmicutes; Bacillales; Oceanobacillus.
NCBI_TaxID=182710;
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN=HTE831 / DSM 14371 / JCM 11309;
RC MEDLINE=22220767; PubMed=12235376;
RT Takami H., Takaki Y., Uchiyama I.;
RA "Genome sequence of Oceanobacillus iheyensis isolated from the Iheya
RT Ridge and its unexpected adaptive capabilities to extreme
RT environments.";
RL Nucleic Acids Res. 30:3927-3935(2002).
RW EMBL; AP004604; BAC15339.1; --
KW Hypothetical protein; Complete proteome.
SQ SEQUENCE 73 AA; 8225 MW; 5BFPE3515EB4A152 CRC64;

Query Match 25.7%; Score 39; DB 16; Length 73;
Best Local Similarity 27.6%; Pred. No. 3.5e+02;
Matches 8; Conservative 9; Mismatches 8; Indels 4; Gaps 0;

Qy 1 YDALNVILVAMNIISKEKEIKWIGLPTNS 29
Db 5 YATWGILTGIFIIA---FLRWIGVPISA 29

RESULT 26
Q8CZ72 PRELIMINARY; PRT; 80 AA.
AC AC
Q8CZ72; 2003 (TrEMBLrel. 23, Created)
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE Hypothetical protein.
DE Hypothetical protein.
DE SPR0300.
OS Streptococcus pneumoniae (strain ATCC BAA-255 / R6).
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC Streptococcus.
OX NCBI_TaxID=171101;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=21429245; PubMed=11544234;
RA Hoskins J., Alborn W.E. Jr., Arnold J., Blaszcak L.C., Burgett S.,
RA Dehoff B.S., Estrem S.T., Fritz L., Fu D.-J., Fuller W., Geringer C.,
RA Gilmour R., Glass J.S., Khoja H., Kraft A.R., Lagace R.E.,
RA Leblanc D.J., Lee L.N., Lefkowitz E.J., Lu J., Matsushima P.,
RA McAhren S.H., McHenry M., McLeaster K., Mundy C.W., Nicas T.I.,
RA Norris F.H., O'Gara M., Peery R.B., Robertson G.T., Rockey P.,
RA Sun P.-M., Winkler M.E., Yang Y., Young-Bellido M., Zhao G.,
RA Zook C.A., Baltz R.H., Jaskunas S.R., Rostek P.R. Jr., Skatrud P.L.,
RA Glass J.I.;
RT "Genome of the bacterium Streptococcus pneumoniae strain R6.";
RL J. Bacteriol. 183:5709-5717(2001).
RW EMBL; AE008411; AAK99104.1; --
KW Hypothetical protein; Complete proteome.
SQ SEQUENCE 80 AA; 9100 MW; 72BA09DA583E8990 CRC64;

Query Match 25.7%; Score 39; DB 16; Length 80;
Best Local Similarity 41.2%; Pred. No. 3.9e+02;
Matches 7; Conservative 5; Mismatches 5; Indels 0; Gaps 0;

Qy 13 ISKEKEIKWIGLPTNS 29
Db 1 MKKEKQLRYPLKAGS 17

RESULT 27
Q95JP8 PRELIMINARY; PRT; 83 AA.
AC AC
Q95JP8;
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)

```

```

AC Q9K1V1;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DE 01-OCT-2001 (TrEMBLrel. 18, Last annotation update)
DE Hypothetical protein CP0930.
GN CP0930.
OS Chlamydia pneumoniae (Chlamydia pneumoniae).
OC Bacteria; Chlamydiae; Chlamydiales; Chlamydiaceae; Chlamydia.
OX NCBI_TaxID=83559;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=AR39;
RX MEDLINE=20150255; PubMed=10684935;
RA Read T.D., Brunham R.C., Shen C., Gill S.R., Heidelberg J.P.,
RA White O., Hickey E.K., Peterson J., Utterback T., Berry K., Bass S.,
RA Linher K., Weidman J., Khouri H., Craven B., Bowman C., Dodson R.,
RA Gwin M., Nelson W., DeBoy R., Kolonay J., McClarty G., Salzberg S.L.,
RA Eisen J., Fraser C.M.;
RA "Genome sequences of Chlamydia trachomatis MoPn and Chlamydia
RT pneumoniae AK39.";
RL Nucleic Acids Res. 28:1397-1406(2000).
DR EMBL; AB002251; AAF38713.1; -.
DR TIGR; CP0930; -.
KW Hypothetical protein.
SQ SEQUENCE 45 AA; 5716 MW; 499DD7697EE4BE3 CRC64;

Query Match 25.3%; Score 38.5; DB 16; Length 45;
Best Local Similarity 43.5%; Pred. No. 2.6e+02;
Matches 10; Conservative 4; Mismatches 8; Indels 1; Gaps 1;

Qy 1 YDALNVLMMNNIISKEKEIKWI 23
| : | : | : | : | : | : | :
Db 4 YEILRIEMRP-FISFEKKEICYL 25

RESULT 30
Q9USG3
ID Q9USG3 PRELIMINARY; PRT; 31 AA.
AC Q9USG3;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE Hypothetical Trp-Asp repeats containing protein (fragment).
GN SPAC18B11.10.
OS Schizosaccharomyces pombe (Fission yeast).
OC Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;
OC Schizosaccharomycetales; Schizosaccharomycetaceae;
OC Schizosaccharomycetes.
OX NCBI_TaxID=4896;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=968 h90;
RX MEDLINE=20223868; PubMed=10759889;
RA Ding D.Q., Tomita Y., Yamamoto A., Chikashige Y., Haraguchi T.,
RA Hiraoka Y.;
RT "Large-scale screening of intracellular protein localization in living
RT fission yeast cells by the use of a GFP-fusion genomic DNA library.";
RL Genes Cells 5:169-190(2000).
DR EMBL; AB027768; BAA87072.1; -.
FT NON_TER 1
FT NON_TER 31
SQ SEQUENCE 31 AA; 3669 MW; 1523469FBAE7E3B7 CRC64;

Query Match 25.0%; Score 38; DB 3; Length 31;
Best Local Similarity 38.9%; Pred. No. 2.1e+02;
Matches 7; Conservative 7; Mismatches 4; Indels 0; Gaps 0;

Qy 13 ISKEKEIKWIGLPTNSA 30
| : | : | : | : | : | :
Db 5 INKLRELEUGVEANTA 22

RESULT 31

```

```

Q8KYL2
ID Q8KYL2 PRELIMINARY; PRT; 39 AA.
AC Q8KYL2;
DT 01-OCT-2002 (TrEMBLrel. 22, Created)
DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
DT 01-OCT-2002 (TrEMBLrel. 22, Last annotation update)
DE Hypothetical protein.
GN BXA0150.
OS Bacillus anthracis.
OX NCBI_TaxID=1392;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=A2012;
RX MEDLINE=22061436; PubMed=12004073;
RA Read T.D., Salzberg S.L., Pop M., Shumway M., Umayam L., Jiang L.,
RA Holtzapfle E., Busch J.D., Smith K.L., Schupp J.M., Solomon D.,
RA Keim P., Fraser C.W.;
RA "Comparative Genome Sequencing for Discovery of Novel Polymorphisms in
RT Bacillus anthracis.";
RL Science 296:2028-2033(2002).
DR EMBL; AE011190; AAM26095.1; -.
KW Hypothetical protein; Plasmid.
SQ SEQUENCE 39 AA; 4581 MW; 8D485BCDCCE3E177 CRC64;

Query Match 25.0%; Score 38; DB 2; Length 39;
Best Local Similarity 37.5%; Pred. No. 2.7e+02;
Matches 6; Conservative 4; Mismatches 6; Indels 0; Gaps 0;

Qy 10 MNIIISKEKEIKWIGL 25
| : | : | : | : | : | :
Db 1 MDIFRKNEKSILWVNM 16

RESULT 32
Q8NVK6
ID Q8NVK6 PRELIMINARY; PRT; 44 AA.
AC Q8NVK6;
DT 01-OCT-2002 (TrEMBLrel. 22, Created)
DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
DT 01-OCT-2002 (TrEMBLrel. 22, Last annotation update)
DE Delta-hemolysin.
GN HLD OR MW1959.
OS Staphylococcus aureus (strain MW2).
OC Bacteria; Firmicutes; Bacillales; Staphylococcus.
OX NCBI_TaxID=196620;
RN [1]
RP SEQUENCE FROM N.A.
RC MEDLINE=22040717; PubMed=12044378;
RA Baba T., Takeuchi F., Kuroda M., Yuzawa H., Aoki K.-I., Oguchi A.,
RA Nagai Y., Iwama N., Asano K., Naimi T., Kuroda H., Cui L.,
RA Yamamoto K., Hiramatsu K.;
RT "Genome and virulence determinants of high virulence community-
RT acquired MRSA.";
RL Lancet 359:1819-1827(2002).
DR EMBL; AP004829; BAB95824.1; -.
KW Complete proteome.
SQ SEQUENCE 44 AA; 5039 MW; 24B1A61E598DF6A8 CRC64;

Query Match 25.0%; Score 38; DB 16; Length 44;
Best Local Similarity 44.4%; Pred. No. 3e+02;
Matches 8; Conservative 3; Mismatches 7; Indels 0; Gaps 0;

Qy 6 VLMAMNIIISKEKEIKWI 23
| : | : | : | : | : | :
Db 17 ISMAQDIISTISDLVKWI 34

RESULT 33
Q98P05
ID Q98P05 PRELIMINARY; PRT; 48 AA.
AC Q98P05;

```

```

DT 01-OCT-2001 (TReMBLrel. 18, Created)
DT 01-OCT-2001 (TReMBLrel. 18, Last sequence update)
DT 01-MAR-2002 (TReMBLrel. 20, Last annotation update)
DE Hypothetical protein msr9763.
GN MSR9763.
OS Rhizobium loti (Mesorhizobium loti).
ON
OG Bacterioidetes; Alphaproteobacteria; Rhizobiales;
OC Bacterioidetes; Alphaproteobacteria; Rhizobiales;
OC Phyllobacteriaceae; Mesorhizobium.
OX NCBI_TaxID=381;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=MAFF303099; PubMed=11214968;
RX MEDLINE=21082930; Sato S., Asamizu E., Kato T., Sasamoto S.,
RA Kaneko T., Nakamura Y., Sato S., Idegawa K., Ishikawa A., Kawashima K., Kimura T.,
RA Watanabe A., Idegawa K., Idegawa K., Kohara M., Matsumoto M., Matsuno A.,
RA Kishida Y., Kiyokawa C., Kohara M., Matsumoto M., Matsuno A.,
RA Mochizuki Y., Nakayama S., Nakazaki N., Shimpo S., Sugimoto M.,
RA Takeuchi C., Yamada M., Tabata S.;
RT "Complete genome structure of the nitrogen-fixing symbiotic bacterium
RT Mesorhizobium loti.";
RT DNA Res. 7:331-338(2000).
DR EMBL; AP003017; BAB54850.1; --
KW Plasmid; Hypothetical protein; Complete proteome.
SQ SEQUENCE 48 AA; 5187 MW; ABA2FC02CA77D6C4 CRC64;

Query Match 25.0%; Score 38; DB 16; Length 48;
Best Local Similarity 40.0%; Pred. No. 3.3e+02;
Matches 8; Conservative 7; Mismatches 3; Indels 2; Gaps 1;

QY 10 MNIISKEKEIKWIGLPTNS 29
Db 1 MNLIS--KGQLEVGLEPAHA 18

RESULT 34
ID Q8RYN3 PRELIMINARY; PRT; 62 AA.
AC Q8RYN3
DT 01-JUN-2002 (TReMBLrel. 21, Created)
DT 01-JUN-2002 (TReMBLrel. 21, Last sequence update)
DT 01-OCT-2002 (TReMBLrel. 22, Last annotation update)
DE P0663E10.12 protein.
GN P0663E10.12
OS Oryza sativa (japonica cultivar-group).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC Ehrhartoideae; Oryzaceae; Oryza.
OX NCBI_TaxID=39947;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=cv. Nipponbare;
RA Sasaki T., Matsumoto T., Yamamoto K.;
RA "Oryza sativa nipponbare(GA3) genomic DNA, chromosome 1, PAC
RT clone.P0663E10.";
RL Submitted (NOV-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AP004317; BAB90729.1; --
DR Gramineae; Q8RYN3; --
SQ SEQUENCE 62 AA; 7432 MW; 8A2F84AD87B913EA CRC64;

Query Match 25.0%; Score 38; DB 10; Length 62;
Best Local Similarity 58.3%; Pred. No. 4.2e+02;
Matches 7; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 15 KEKKEIKWIGLP 26
Db 32 KRKEIEWIENP 43

RESULT 35
Q9JF42 PRELIMINARY; PRT; 76 AA.
ID Q9JF42
AC Q9JF42

```

```

DT 01-OCT-2000 (TReMBLrel. 15, Created)
DT 01-OCT-2000 (TReMBLrel. 15, Last sequence update)
DT 01-OCT-2000 (TReMBLrel. 15, Last annotation update)
DE TB1LR.
OS Vaccinia virus (strain Tian Tan).
OC Viruses; dsDNA viruses, no RNA stage; Poxviridae; Chordopoxvirinae;
OC Orthopoxvirus.
OX NCBI_TaxID=10253;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Tian Tan;
RA Jin Q., Hou Y.D., Cheng N.H., Yao E.M., Cheng S.X., Yang X.K.,
RA Jing D.Y., Yu W.H., Yuan J.S., Ma X.J.;
RT "Complete genomic sequence of vaccinia virus (Tian Tan strain).";
RL Submitted (SEP-1998) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF095689; AAF34080.1; --
SQ SEQUENCE 76 AA; 8616 MW; 7066B8F1A9D2F522 CRC64;

Query Match 25.0%; Score 38; DB 12; Length 76;
Best Local Similarity 47.4%; Pred. No. 5.2e+02;
Matches 9; Conservative 3; Mismatches 7; Indels 0; Gaps 0;

QY 2 DALNVLMMNIIISKEKEI 20
Db 8 DVTNVEDIINEIDREKEI 26

RESULT 36
ID Q9B911 PRELIMINARY; PRT; 81 AA.
AC Q9B911
DT 01-JUN-2001 (TReMBLrel. 17, Created)
DT 01-JUN-2001 (TReMBLrel. 17, Last sequence update)
DT 01-OCT-2002 (TReMBLrel. 22, Last annotation update)
DE Cytochrome oxidase subunit II (COII) gene (Fragment).
OS Wiebesia brusi.
OG Mitochondrion.
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Hymenoptera; Apocrita; Chalcidoidea;
OC Agaonidae; Agaoninae; Wiebesia.
OX NCBI_TaxID=130036;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=B120;
RA Weiblen G.D.;
RT "Phylogenetic analyses of dioecious fig pollinators based on
RT mitochondrial DNA sequences and morphology.";
RL Submitted (OCT-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF200412; AAK00126.1; --
DR InterPro; IPR002429; Cyt_c_ox_2.
DR Pfam; PF02790; COX2_TN; I_
DR PRINTS; PR01166; CYCOXIDASEII.
DR Oxidoreductase; Transmembrane; Mitochondrion.
KW NON TER 81
FT SEQUENCE 81 AA; 9471 MW; 921D08E561CAB91A CRC64;

Query Match 25.0%; Score 38; DB 8; Length 81;
Best Local Similarity 36.7%; Pred. No. 5.5e+02;
Matches 11; Conservative 5; Mismatches 10; Indels 4; Gaps 1;

QY 1 YDALNVLMMNII-----ISKEKEIKWIGLP 26
Db 40 YVLLNMLINKNIDRMISGQMIELVMTVVP 69

RESULT 37
ID Q8Y9B0 PRELIMINARY; PRT; 92 AA.
AC Q8Y9B0
DT 01-MAR-2002 (TReMBLrel. 20, Created)
DT 01-MAR-2002 (TReMBLrel. 20, Last sequence update)
DT 01-MAR-2002 (TReMBLrel. 20, Last annotation update)
DE Hypothetical protein lmo0622.

```



```
GN LMO0622.
OS Listeria monocytogenes.
OC Bacteria; Firmicutes; Bacillales; Listeriaceae; Listeria.
OX NCBI_TaxID=1639;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=EGD-e / Serovar 1/2a;
RX MEDLINE=21537279; PubMed=11679669;
RA Glaser P., Frangeul L., Buchrieser C., Rusniok C., Amend A.,
RA Baquero F., Berche P., Bloeker H., Brandt P., Chakraborty T.,
RA Charbit A., Cherouani F., Couve E., de Daruvar A., Dehoux P.,
RA Domann E., Dominguez-Bernal G., Duchaud E., Durant L., Dussurget O.,
RA Etian K.-D., Faihi H., Garcia-del Portillo F., Garrido P.,
RA Gautier L., Goebel W., Gomez-Lopez N., Hain T., Hauf J., Jackson D.,
RA Jones L.-M., Kaefer U., Kreft J., Kuhn M., Kunst F., Kurupkat G.,
RA Madueno E., Maitournam A., Mata Vicente J., Ng E., Nedjari H.,
RA Nordsiek G., Novella S., de Pablo B., Perez-Diaz J.-C., Purcell R.,
RA Rammel B., Rose M., Schluter T., Simoes N., Tierrez A.,
RA Vazquez-Boland J.-A., Voss H., Wehland J., Cossart P.;
RT "Comparative genomics of Listeria species.";
RL Science 294:849-852(2001).
DR EMBL; AL591976; CAC98700.1; .
DR L1000; LMO0622; .
KW Hypothetical protein; Complete proteome.
SQ SEQUENCE 92 AA; 10656 MW; B1712506710528FA CRC64;

Query Match 25.0%; Score 38; DB 16; Length 92;
Best Local Similarity 28.0%; Pred. No. 6.3e+02;
Matches 7; Conservative 8; Mismatches 10; Indels 0; Gaps 0;

Qy 4 LNVLMANNIISKEKEIKWIGLPTN 28
Db 49 LGIISTISVFFVREKLVKILLGFN 73

RESULT 38
Q9BQR8
ID Q9BQR8 PRELIMINARY; PRT; 95 AA.
AC Q9BQR8;
DT 01-JUN-2001 (TrEMBLrel. 17, Created)
DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE Chr14 synaptotagmin (fragment).
GN CHR14SYT.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Brain;
RX MEDLINE=21429112; PubMed=11543631;
RA Craxton M.;
RT "Genomic analysis of synaptotagmins.";
RL Genomics 77:43-49(2001).
DR EMBL; AJ303367; CAC33889.1; .
DR InterPro; IPR000008; C2.
DR InterPro; IPR002149; LRI.
DR Pfam; PF00168; C2; 1.
FT NON_TER 1 95
FT NON_TER 95 95
SQ SEQUENCE 95 AA; 11168 MW; CD748C4CB5643A CRC64;

Query Match 25.0%; Score 38; DB 4; Length 95;
Best Local Similarity 50.0%; Pred. No. 6.5e+02;
Matches 8; Conservative 2; Mismatches 6; Indels 0; Gaps 0;

Qy 15 KEKEIKWIGLPTNSA 30
Db 52 KRKEMIGWALQNS 67

RESULT 39
Q9B3E7
ID Q9B3E7 PRELIMINARY; PRT; 96 AA.
AC Q9B3E7;
DT 01-MAR-2002 (TrEMBLrel. 20, Created)
DT 01-MAR-2002 (TrEMBLrel. 20, Last sequence update)
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE RNA dependent RNA polymerase (fragment).
GN RDRP.
OS Apricot latent virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage.
OX NCBI_TaxID=75387;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Caserta12;
RA Poiseac X., Svanella-Dumas L., Gentit P., Dulucq M.-J., Candresse T.;
RT "PDO RT-PCR : a polyvalent detection tool for an etiological and
RT variability studies for tricho-, capillo- and foveaviruses of
RT Prunus.";
RL Submitted (AUG-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF413914; AAL60487.1; .
DR InterPro; IPR001788; RNA_dep_RNApol2.
DR Pfam; PF00978; RNA_dep_RNApol2; 1.
FT NON_TER 1 96
FT NON_TER 96 96
SQ SEQUENCE 96 AA; 10982 MW; E264C91952F00474 CRC64;

Query Match 25.0%; Score 38; DB 12; Length 96;
Best Local Similarity 28.6%; Pred. No. 6.5e+02;
Matches 8; Conservative 9; Mismatches 11; Indels 0; Gaps 0;

Qy 1 YDALNVLMMNIIISKEKEIKWIGLPTN 28
Db 50 YEAFDASQDHFILAFELVNMKFLGLPAD 77

RESULT 40
Q8V3E1
ID Q8V3E1 PRELIMINARY; PRT; 96 AA.
AC Q8V3E1;
DT 01-MAR-2002 (TrEMBLrel. 20, Created)
DT 01-MAR-2002 (TrEMBLrel. 20, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE RNA dependent RNA polymerase (fragment).
GN RDRP.
OS Apricot latent virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage.
OX NCBI_TaxID=75387;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ALV;
RA Poiseac X., Svanella-Dumas L., Gentit P., Dulucq M.-J., Candresse T.;
RT "PDO RT-PCR : a polyvalent detection tool for an etiological and
RT variability studies for tricho-, capillo- and foveaviruses of
RT Prunus.";
RL Submitted (AUG-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF413920; AAL60493.1; .
DR InterPro; IPR001788; RNA_dep_RNApol2.
DR InterPro; IPR007094; RNA_pol_PSVir.
DR Pfam; PF00978; RNA_dep_RNApol2; 1.
DR PROSITE; PSS0507; RDRP_POSITIVE; 1.
FT NON_TER 1 96
FT NON_TER 96 96
SQ SEQUENCE 96 AA; 11189 MW; 4BDBFDE2567ECB3D CRC64;

Query Match 25.0%; Score 38; DB 12; Length 96;
Best Local Similarity 28.6%; Pred. No. 6.5e+02;
Matches 8; Conservative 9; Mismatches 11; Indels 0; Gaps 0;

Qy 1 YDALNVLMMNIIISKEKEIKWIGLPTN 28
Db 50 YEAFDASQDHFILAFELVNMKFLGLPAD 77
```

Tue Feb 17 11:55:55 2004

```

RESULT 41
Q8RF87 PRELIMINARY; PRT; 97 AA.
AC Q8RF87;
ID Q8RF87;
DT 01-JUN-2002 (TrEMBLrel. 21, Created)
DT 01-JUN-2002 (TrEMBLrel. 21, Last sequence update)
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE Hypothetical protein FN0829.
DE FN0829.
GN FUSOBACTERIUM NUCLEATUM (subsp. nucleatum).
OS FUSOBACTERIA; FUSOBACTERIA; FUSOBACTERIALES; FUSOBACTERIACEAE;
OC Bacteria;
OC FUSOBACTERIUM.
OC NCBI_TaxID=76856;
OX (1)
RN SEQUENCE FROM N.A.
RP STRAIN=ATCC 25586;
RC MEDLINE=21886394; PubMed=11889109;
RX Kapral V., Anderson I., Ivanova N., Reznik G., Los T., Lykidis A.,
RA Bhattacharyya A., Bartman A., Gardner W., Grechkin G., Zhu L.,
RA Vasileva O., Chu L., Kogan Y., Chaga O., Goltsman E., Bernal A.,
RA Larsen N., D'Souza M., Walunas T., Pusch G., Haselkorn R.,
RA Fongstein M., Kyrpides N., Overbeek R.;
RT "Genome sequence and analysis of the oral bacterium Fusobacterium
nucleatum strain ATCC 25586."
RL J. Bacteriol. 184:2005-2018(2002).
RL EMBL; AF010592; AAL95025.1; -.
KW Hypothetical protein; Complete proteome.
SQ SEQUENCE 97 AA; 11914 MW; D16C48E41F9FD7EA CRC64;

Query Match 25.0%; Score 38; DB 16; Length 97;
Best Local Similarity 43.8%; Pred. No. 6.6e+02;
Matches 7; Conservative 7; Mismatches 2; Indels 0; Gaps 0;

QY 4 LNVLMAMNIIISKEKE 19
Db 8 LLLILSISIFSQEKKE 23

RESULT 42
Q8CSCO PRELIMINARY; PRT; 97 AA.
AC Q8CSCO;
ID Q8CSCO;
DT 01-MAR-2003 (TrEMBLrel. 23, Created)
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE Conserved hypothetical protein.
DE SEI281.
GN Staphylococcus epidermidis.
OS Bacteria; Firmicutes; Bacillales; Staphylococcus.
OC NCBI_TaxID=1282;
OX (1)
RN SEQUENCE FROM N.A.
RP STRAIN=ATCC 12228;
RC Zhang Y., Ren S., Li H., Fu G., Lu L., Lu G., Jia J., Tu Y., Qin Z.,
RA Chen Z., Wen Y.;
RA Submitted (NOV-2002) to the EMBL/GenBank/DBJ databases.
RL EMBL; AE016748; AAO04880.1; -.
KW Hypothetical protein; Complete proteome.
SQ SEQUENCE 97 AA; 11165 MW; 8FB484B1968C4ECC CRC64;

Query Match 25.0%; Score 38; DB 16; Length 97;
Best Local Similarity 32.0%; Pred. No. 6.6e+02;
Matches 8; Conservative 8; Mismatches 9; Indels 0; Gaps 0;

QY 2 DALNVLAMNIIISKEKEIKWIGLP 26
Db 72 ELVQIIGSMIVLYKESEENKEIKLP 96

RESULT 43
Q9YI44 PRELIMINARY; PRT; 68 AA.
ID Q9YI44;
AC Q9YI44;

RESULT 44
Q9K761 PRELIMINARY; PRT; 83 AA.
AC Q9K761;
ID Q9K761;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-MAR-2002 (TrEMBLrel. 20, Last annotation update)
DE Hypothetical protein BH3512.
GN BH3512.
OS Bacillus halodurans.
OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.
OC NCBI_TaxID=86665;
OX (1)
RN SEQUENCE FROM N.A.
RP STRAIN=C-125 / JCM 9153;
RC MEDLINE=20512582; PubMed=11058132;
RX Takami H., Nakasone K., Takaki Y., Maeno G., Sasaki R., Masui N.,
RA Fuji F., Hirama C., Nakamura Y., Ogasawara N., Kuhara S.,
RA Horikoshi K.;
RT "Complete genome sequence of the alkaliphilic bacterium Bacillus
halodurans and genomic sequence comparison with Bacillus subtilis."
RL Nucleic Acids Res. 28:4317-4331(2000).
RL EMBL; AP001519; BAB07231.1; -.
KW Hypothetical protein; Complete proteome.
SQ SEQUENCE 83 AA; 9454 MW; 471552D292B2A44F CRC64;

Query Match 24.7%; Score 37.5; DB 16; Length 83;
Best Local Similarity 27.7%; Pred. No. 6.7e+02;
Matches 13; Conservative 7; Mismatches 10; Indels 17; Gaps 2;

QY 1 YDALNVLMA-----MNIISKEKEIK--WIGLPTNSA 30
Db 17 YDTLGLDSSVKRLQRIIRLEKTSLAEEKREIKKWLAFLSA 63

RESULT 45

```

Search completed: February 17, 2004, 10:56:24  
Job time : 44.4752 secs

**THIS PAGE BLANK (USPTO)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: February 17, 2004, 10:50:13 ; Search time 9.50495 Seconds  
(without alignments)  
148.428 Million cell updates/sec

Title: US-09-900-147-6  
Perfect score: 152  
Sequence: 1 YDALNVLMMANNIISKEKKEIKWIGLPTNSA 30

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 127863 seqs, 47026705 residues

Total number of hits satisfying chosen parameters: 13973

Minimum DB seq length: 0  
Maximum DB seq length: 100

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : SwissProt\_41.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	44	28.9	65	V089_FOWPV	072899 fowlpox vir
2	39	25.7	64	YC53_ARCFU	Q29015 archaeoglob
3	38.5	25.3	99	Y246_METJA	Q57696 methanococ
4	38	25.0	72	YB11_VACCV	Q01229 vaccinia vi
5	38	25.0	88	YB11_VACCV	P21007 vaccinia vi
6	38	25.0	91	VAPD_HAEIN	P71351 haemophilus
7	37	24.3	87	BOFA_BACSU	P24282 bacillus su
8	37	24.3	88	RL29_SULTO	Q97518 sulfobolus
9	36	23.7	45	HLD_STAAM	P01506 staphylococ
10	36	23.7	56	Y546_METJA	Q57966 methanococ
11	36	23.7	74	Y295_ARCFU	Q29947 archaeoglob
12	36	23.7	88	YDFK_ECOLI	P76154 escherichia
13	36	23.7	88	YNAE_ECOLI	P76073 escherichia
14	36	23.7	96	YQEI_BACSU	P54454 bacillus su
15	36	23.7	97	YHBY_ECOLI	P42550 escherichia
16	35	23.0	68	TRPC_BACPU	P18268 bacillus pu
17	35	23.0	72	COX2_LEPSP	P29659 lepisosteus
18	35	23.0	72	COX2_LEPSP	P29656 lepisosteus
19	35	23.0	86	FCEG_MOUSE	P20491 mus musculus
20	35	23.0	95	Y511_BUCAP	Q8K946 buchnera ap
21	35	23.0	98	RS24_THEAC	Q9HJ79 thermoplas
22	34.5	22.7	52	NU3M_NEUCR	Q35141 neurospora
23	34.5	22.7	90	RS16_STR3A	Q8e4h3 streptococ
24	34	22.4	59	SECE_BACLI	P38381 bacillus li
25	34	22.4	62	YM32_MARPO	P38474 marchantia
26	34	22.4	70	CSPA_RICCN	Q92Gv1 rickettsia
27	34	22.4	70	CSPA_RICPR	Q92Cp9 rickettsia
28	34	22.4	72	RL15_BACLI	P35138 bacillus li
29	34	22.4	84	LEU2_HUMAN	O43262 homo sapien
30	34	22.4	86	FCEG_RAT	P20411 rattus norv
31	34	22.4	90	Y530_BUCAI	P57596 buchnera ap
32	34	22.4	95	RL14_CARRU	Q9aif4 carsonella
33	34	22.4	95	Y117_NPVAC	P41670 autographa

34 34 22.4 97 1 SPAC\_BPT4 P39230 bacterioph  
35 33.5 22.0 42 1 RET5\_BOVIN P82708 bos taurus  
36 33.5 22.0 77 1 ACPH\_UREPA Q9ppv4 ureaplasma  
37 33.5 22.0 90 1 RS16\_STRPY P58124 streptococ  
38 33.5 22.0 91 1 RS16\_STRMU Q8dun9 streptococ  
39 33.5 22.0 94 1 RUXE\_YEAST Q12330 saccharomyc  
40 33 21.7 42 1 RL32\_MAIZE P51421 zea mays (m  
41 33 21.7 59 1 SECE\_BACSU Q06799 bacillus su  
42 33 21.7 59 1 YDBF\_SCHPO Q10365 schizosacch  
43 33 21.7 64 1 GEPD\_BACCR O68686 bacillus ce  
44 33 21.7 65 1 Y12J\_BPT4 Q02406 bacterioph  
45 33 21.7 83 1 VG5\_SPV4 P11337 spiropasma

ALIGNMENTS

RESULT 1  
V089\_FOWPV STANDARD; PRT; 65 AA.  
ID V089\_FOWPV  
AC 072899;  
DT 16-OCT-2001 (Rel. 40, Created)  
DT 16-OCT-2001 (Rel. 40, Last sequence update)  
DT 16-OCT-2001 (Rel. 40, Last annotation update)  
DE Protein FPV089.  
GN FPV089 OR FP12L.  
OS Fowlpox virus (FPV).  
OC Viruses; dsDNA viruses, no RNA stage; Poxviridae; Chordopoxvirinae;  
OC Avipoxvirus.  
OC NCBI\_TaxID=10261;  
OX [1]  
RN RP  
RP SEQUENCE FROM N.A.  
RC STRAIN=FP-9 / Isolate HP-440;  
RX MEDLINE=98451804; PubMed=9778782;  
RA Pollitt E., Skinner M.A., Heaphy S.;  
RT "Nucleotide sequence of the 4.3 kbp BamHI-N fragment of fowlpox virus  
FP9.";  
RL Virus Genes 17:5-9 (1998).  
RN [2]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=20193820; PubMed=10729156;  
RA Alfonso C.L., Tulman E.R., Lu Z., Zsak L., Kutish G.F., Rock D.L.;  
RT "The genome of fowlpox virus.";  
RL J. Virol. 74:3815-3831 (2000).  
CC -|- SIMILARITY: BELONGS TO THE POXVIRUSES I2 FAMILY.  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
DR EMBL; AJ223385; CAAL1292.1; -  
DR EMBL; AF198100; AAF4433.1; -  
SQ SEQUENCE 65 AA; 7573 MW; ECDD9A25B824DF92 CRC64;  
Query Match 28.9%; Score 44; DB 1; Length 65;  
Best Local Similarity 31.8%; Pred. No. 8.1;  
Matches 7; Conservative 6; Mismatches 9; Indels 0; Gaps 0;  
QY 2 DALNVLMMANNIISKEKKEIKWI 23  
Db 22 DFINTIMVLTKKESKQLSWL 43  
RESULT 2  
YC53\_ARCFU STANDARD; PRT; 64 AA.  
ID YC53\_ARCFU  
AC O29015;  
DT 16-OCT-2001 (Rel. 40, Created)  
DT 16-OCT-2001 (Rel. 40, Last sequence update)

16-OCT-2001 (Rel. 40, Last annotation update)  
Hypothetical protein AF1253.  
Archaeoglobus fulgidus.  
Archaea; Euryarchaeota; Archaeoglobales;  
Archaeoglobaceae; Archaeoglobus.  
NCBI\_TaxID=2234;  
[1]  
SEQUENCE FROM N.A.  
STRAIN=VC-16 / DSM 4304 / ATCC 49558;  
MEDLINE=98049343; PubMed=9389475;  
Klenk H.-P., Clayton R.A., Tomb J.-F., White O., Nelson K.E.,  
Ketchum K.A., Dodson R.J., Winn M., Hickey E.K., Peterson J.D.,  
Richardson D.L., Keriavag A.J., Graham D.E., Kyprides N.C.,  
Fleischmann R.D., Quackenbush J., Lee N.H., Sutton G.G., Gill S.,  
Kirkness E.F., Dougherty B.A., McKenney K., Adams M.D., Loftus B.,  
Peterson S., Reich C.I., McNeill L.K., Badger J.H., Glodek A., Zhou L.,  
Overbeek R., Gocayne J.D., Weidman J.F., McDonald L., Uterback T.,  
Cotton M.D., Spriggs T., Artach P., Kaine B.P., Sykes S.M.,  
Sadow F.W., D'Andrea K.P., Bowman C., Fujii C., Garland S.A.,  
Mason T.M., Olsen G.J., Fraser C.M., Smith H.O., Woese C.R.,  
Venter J.C.;  
"The complete genome sequence of the hyperthermophilic, sulphate-  
reducing archaeon Archaeoglobus fulgidus.";  
Nature 390:364-370(1997).  
-----  
This SWISS-PROT entry is copyright. It is produced through a collaboration  
between the Swiss Institute of Bioinformatics and the EMBL outstation -  
the European Bioinformatics Institute. There are no restrictions on its  
use by non-profit institutions as long as its content is in no way  
modified and this statement is not removed. Usage by and for commercial  
entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>  
or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch))  
-----  
EMBL; AF001018; AAB90006.1; -;  
PIR; D69406; D69406.  
TIGR; AF1253; -;  
Hypothetical protein; Complete proteome.  
SEQUENCE 64 AA; 7418 MW; CD6CF9F885F4FF15 CRC64;  
-----  
Query Match 25.7%; Score 39; DB 1; Length 64;  
Best Local Similarity 50.0%; Pred. No. 45; Indels 0; Gaps 0;  
Matches 10; Conservative 4; Mismatches 5;  
-----  
QY 4 LNVLMAMNIISKEKEIKWI 23  
DB 42 LDKIEAVEITRAKKEIKI 61  
-----  
RESULT 3  
Y246 METJA STANDARD; PRT; 99 AA.  
ID Y246 METJA STANDARD; PRT;  
AC 057656;  
DT 01-NOV-1997 (Rel. 35, Created)  
DT 01-NOV-1997 (Rel. 35, Last sequence update)  
DT 16-OCT-2001 (Rel. 40, Last annotation update)  
DE DE Hypothetical protein M70246.  
GN M70246.  
OS Methanococcus jannaschii.  
OC Archaea; Euryarchaeota; Methanococci; Methanococcales;  
OC Methanocaldococaceae; Methanocaldococcus.  
OX NCBI\_TaxID=2190;  
[1]  
SEQUENCE FROM N.A.  
STRAIN=JAL-1 / DSM 2661 / ATCC 43067;  
MEDLINE=96337999; PubMed=8688087;  
Bult C.J., White O., Olsen G.J., Zhou L., Fleischmann R.D.,  
Sutton G.G., Blake J.A., FitzGerald L.M., Clayton R.A., Gocayne J.D.,  
Keriavag A.R., Dougherty B.A., Tomb J.-F., Adams M.D., Reich C.I.,  
Overbeek R., Kirkness E.F., Weinstock K.G., Merrick J.M., Glodek A.,  
Scott J.L., Geoghagen N.S.M., Weidman J.F., Fuhrmann J.L., Nguyen D.,  
Uterback T.R., Kelley J.M., Peterson J.D., Sadow F.W., Hanna M.C.,  
Cotton M.D., Roberts K.M., Hurst M.A., Kaine B.P., Borodovsky M.,  
Venter J.C.;  
"The complete genome sequence of the hyperthermophilic, sulphate-  
reducing archaeon Archaeoglobus fulgidus.";  
Nature 390:364-370(1997).  
-----  
This SWISS-PROT entry is copyright. It is produced through a collaboration  
between the Swiss Institute of Bioinformatics and the EMBL outstation -  
the European Bioinformatics Institute. There are no restrictions on its  
use by non-profit institutions as long as its content is in no way  
modified and this statement is not removed. Usage by and for commercial  
entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>  
or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch))  
-----  
EMBL; AF001018; AAB90006.1; -;  
PIR; D69406; D69406.  
TIGR; AF1253; -;  
Hypothetical protein; Complete proteome.  
SEQUENCE 64 AA; 7418 MW; CD6CF9F885F4FF15 CRC64;  
-----  
Query Match 25.7%; Score 39; DB 1; Length 64;  
Best Local Similarity 50.0%; Pred. No. 45; Indels 0; Gaps 0;  
Matches 10; Conservative 4; Mismatches 5;  
-----  
QY 4 LNVLMAMNIISKEKEIKWI 23  
DB 42 LDKIEAVEITRAKKEIKI 61  
-----  
RESULT 3  
Y246 METJA STANDARD; PRT; 99 AA.  
ID Y246 METJA STANDARD; PRT;  
AC 057656;  
DT 01-NOV-1997 (Rel. 35, Created)  
DT 01-NOV-1997 (Rel. 35, Last sequence update)  
DT 16-OCT-2001 (Rel. 40, Last annotation update)  
DE DE Hypothetical protein M70246.  
GN M70246.  
OS Methanococcus jannaschii.  
OC Archaea; Euryarchaeota; Methanococci; Methanococcales;  
OC Methanocaldococaceae; Methanocaldococcus.  
OX NCBI\_TaxID=2190;  
[1]  
SEQUENCE FROM N.A.  
STRAIN=JAL-1 / DSM 2661 / ATCC 43067;  
MEDLINE=96337999; PubMed=8688087;  
Bult C.J., White O., Olsen G.J., Zhou L., Fleischmann R.D.,  
Sutton G.G., Blake J.A., FitzGerald L.M., Clayton R.A., Gocayne J.D.,  
Keriavag A.R., Dougherty B.A., Tomb J.-F., Adams M.D., Reich C.I.,  
Overbeek R., Kirkness E.F., Weinstock K.G., Merrick J.M., Glodek A.,  
Scott J.L., Geoghagen N.S.M., Weidman J.F., Fuhrmann J.L., Nguyen D.,  
Uterback T.R., Kelley J.M., Peterson J.D., Sadow F.W., Hanna M.C.,  
Cotton M.D., Roberts K.M., Hurst M.A., Kaine B.P., Borodovsky M.,  
Venter J.C.;  
"The complete genome sequence of the hyperthermophilic, sulphate-  
reducing archaeon Archaeoglobus fulgidus.";  
Nature 390:364-370(1997).  
-----  
This SWISS-PROT entry is copyright. It is produced through a collaboration  
between the Swiss Institute of Bioinformatics and the EMBL outstation -  
the European Bioinformatics Institute. There are no restrictions on its  
use by non-profit institutions as long as its content is in no way  
modified and this statement is not removed. Usage by and for commercial  
entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>  
or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch))  
-----  
EMBL; AF001018; AAB90006.1; -;  
PIR; D69406; D69406.  
TIGR; AF1253; -;  
Hypothetical protein; Complete proteome.  
SEQUENCE 64 AA; 7418 MW; CD6CF9F885F4FF15 CRC64;  
-----  
Query Match 25.7%; Score 39; DB 1; Length 64;  
Best Local Similarity 50.0%; Pred. No. 45; Indels 0; Gaps 0;  
Matches 10; Conservative 4; Mismatches 5;  
-----  
QY 4 LNVLMAMNIISKEKEIKWI 23  
DB 42 LDKIEAVEITRAKKEIKI 61  
-----  
RESULT 3  
Y246 METJA STANDARD; PRT; 99 AA.  
ID Y246 METJA STANDARD; PRT;  
AC 057656;  
DT 01-NOV-1997 (Rel. 35, Created)  
DT 01-NOV-1997 (Rel. 35, Last sequence update)  
DT 16-OCT-2001 (Rel. 40, Last annotation update)  
DE DE Hypothetical protein M70246.  
GN M70246.  
OS Methanococcus jannaschii.  
OC Archaea; Euryarchaeota; Methanococci; Methanococcales;  
OC Methanocaldococaceae; Methanocaldococcus.  
OX NCBI\_TaxID=2190;  
[1]  
SEQUENCE FROM N.A.  
STRAIN=JAL-1 / DSM 2661 / ATCC 43067;  
MEDLINE=96337999; PubMed=8688087;  
Bult C.J., White O., Olsen G.J., Zhou L., Fleischmann R.D.,  
Sutton G.G., Blake J.A., FitzGerald L.M., Clayton R.A., Gocayne J.D.,  
Keriavag A.R., Dougherty B.A., Tomb J.-F., Adams M.D., Reich C.I.,  
Overbeek R., Kirkness E.F., Weinstock K.G., Merrick J.M., Glodek A.,  
Scott J.L., Geoghagen N.S.M., Weidman J.F., Fuhrmann J.L., Nguyen D.,  
Uterback T.R., Kelley J.M., Peterson J.D., Sadow F.W., Hanna M.C.,  
Cotton M.D., Roberts K.M., Hurst M.A., Kaine B.P., Borodovsky M.,  
Venter J.C.;  
"The complete genome sequence of the hyperthermophilic, sulphate-  
reducing archaeon Archaeoglobus fulgidus.";  
Nature 390:364-370(1997).  
-----  
This SWISS-PROT entry is copyright. It is produced through a collaboration  
between the Swiss Institute of Bioinformatics and the EMBL outstation -  
the European Bioinformatics Institute. There are no restrictions on its  
use by non-profit institutions as long as its content is in no way  
modified and this statement is not removed. Usage by and for commercial  
entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>  
or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch))  
-----  
EMBL; AF001018; AAB90006.1; -;  
PIR; D69406; D69406.  
TIGR; AF1253; -;  
Hypothetical protein; Complete proteome.  
SEQUENCE 64 AA; 7418 MW; CD6CF9F885F4FF15 CRC64;  
-----  
Query Match 25.7%; Score 39; DB 1; Length 64;  
Best Local Similarity 50.0%; Pred. No. 45; Indels 0; Gaps 0;  
Matches 10; Conservative 4; Mismatches 5;  
-----  
QY 4 LNVLMAMNIISKEKEIKWI 23  
DB 42 LDKIEAVEITRAKKEIKI 61  
-----  
RESULT 3  
Y246 METJA STANDARD; PRT; 99 AA.  
ID Y246 METJA STANDARD; PRT;  
AC 057656;  
DT 01-NOV-1997 (Rel. 35, Created)  
DT 01-NOV-1997 (Rel. 35, Last sequence update)  
DT 16-OCT-2001 (Rel. 40, Last annotation update)  
DE DE Hypothetical protein M70246.  
GN M70246.  
OS Methanococcus jannaschii.  
OC Archaea; Euryarchaeota; Methanococci; Methanococcales;  
OC Methanocaldococaceae; Methanocaldococcus.  
OX NCBI\_TaxID=2190;  
[1]  
SEQUENCE FROM N.A.  
STRAIN=JAL-1 / DSM 2661 / ATCC 43067;  
MEDLINE=96337999; PubMed=8688087;  
Bult C.J., White O., Olsen G.J., Zhou L., Fleischmann R.D.,  
Sutton G.G., Blake J.A., FitzGerald L.M., Clayton R.A., Gocayne J.D.,  
Keriavag A.R., Dougherty B.A., Tomb J.-F., Adams M.D., Reich C.I.,  
Overbeek R., Kirkness E.F., Weinstock K.G., Merrick J.M., Glodek A.,  
Scott J.L., Geoghagen N.S.M., Weidman J.F., Fuhrmann J.L., Nguyen D.,  
Uterback T.R., Kelley J.M., Peterson J.D., Sadow F.W., Hanna M.C.,  
Cotton M.D., Roberts K.M., Hurst M.A., Kaine B.P., Borodovsky M.,  
Venter J.C.;  
"The complete genome sequence of the hyperthermophilic, sulphate-  
reducing

```

RESULT 5
VB11 VACCC STANDARD; PRT; 88 AA.
ID _V11 VACCC STANDARD; PRT; 88 AA.
AC P21007;
DT 01-FEB-1991 (Rel. 17, Created)
DT 01-FEB-1991 (Rel. 17, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Protein B11.
GN B11R.
OS Vaccinia virus (strain Copenhagen).
OC Viruses; dsDNA viruses, no RNA stage; Poxviridae; Chordopoxvirinae;
OC Orthopoxvirus.
OX NCBI_TaxID=10249;
RN [1]
RN SEQUENCE FROM N.A.
RX MEDLINE=91021027; PubMed=22119722;
RA Goebel S.J., Johnson G.P., Perkus M.E., Davis S.W., Winslow J.P.,
RA Paoletti E.;
RT "The complete DNA sequence of vaccinia virus.";
RL Virology 179:247-266(1990).
RN [2]
RN COMPLETE GENOME
RA Goebel S.J., Johnson G.P., Perkus M.E., Davis S.W., Winslow J.P.,
RA Paoletti E.;
RT "Appendix to 'The complete DNA sequence of vaccinia virus'.";
RL Virology 179:517-563(1990).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; M35027; AAA48208.1; -.
DR PIR; A42527; A42527.
KW Repeat.
FT DOMAIN 2 19 9 X 2 AA TANDEM REPEATS OF D-T.
SQ SEQUENCE 88 AA; 9897 MW; F81B3D79229AE02 CRC64;

Query Match 25.0%; Score 38; DB 1; Length 88;
Best Local Similarity 47.4%; Pred. No. 89;
Matches 9; Conservative 3; Mismatches 7; Indels 0; Gaps 0;

QY 2 DALNVLMMNIIISKEKEI 20
Db 20 DVTNVEDIINEIDREKEEI 38

RESULT 6
VAPD HAEIN STANDARD; PRT; 91 AA.
ID _VAPD HAEIN STANDARD; PRT; 91 AA.
AC P71351;
DT 15-JUL-1999 (Rel. 38, Created)
DT 15-JUL-1999 (Rel. 38, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Virulence-associated protein D.
GN VAPD OR H10450.
OS Haemophilus influenzae.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Pasteurellales;
OC Pasteurellaceae; Haemophilus.
OX NCBI_TaxID=727;
RN [1]
RN SEQUENCE FROM N.A.
RX STRAIN=Rd / KW20 / ATCC 51907;
RC MEDLINE=95350630; PubMed=7542800;
RA Fleischmann R.D., Adams M.D., White O., Clayton R.A., Kirkness E.F.,
RA Kerlavage A.R., Bult C.J., Tomb J.-F., Dougherty B.A., Merrick J.M.,
RA McKenney K., Sutton G., Fitzhugh W., Fields C.A., Gocayne J.D.,
RA Scott J.D., Shirley R., Liu L.-I., Glodek A., Kelley J.M.,
RA Weidman J.F., Phillips C.A., Spriggs T., Hedblom E., Cotton M.D.,

```

```

RA Utterback T.R., Hanna M.C., Nguyen D.T., Saudek D.M., Brandon R.C.,
RA Fine L.D., Fritchman J.L., Fuhrmann J.L., Geoghagen N.S.M.,
RA Gnehm C.L., McDonald L.A., Small K.V., Fraser C.M., Smith H.O.,
RA Venter J.C.;
RT "Whole-genome random sequencing and assembly of Haemophilus influenzae
RT Rd.";
RL Science 269:496-512(1995).
RN [2]
RN SEQUENCE OF 76-80.
RX MEDLINE=20137488; PubMed=10675023;
RA Langen H., Takacs B., Evers S., Berndt P., Lahm H.W., Wipf B.,
RA Gray C., Fountoulakis M.;
RT "Two-dimensional map of the proteome of Haemophilus influenzae.";
RL Electrophoresis 21:411-429(2000).
CC -I- SIMILARITY: BELONGS TO THE VAPD FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; U32728; AAC22108.1; -.
DR PIR; C64069; C64069.
DR TIGR; H10450; -.
DR Pfam; PF04605; VapD N; 1.
KW Virulence; Complete proteome.
SQ SEQUENCE 91 AA; 10343 MW; 70B23CDE28E083E7 CRC64;

Query Match 25.0%; Score 38; DB 1; Length 91;
Best Local Similarity 36.4%; Pred. No. 92;
Matches 8; Conservative 4; Mismatches 6; Indels 4; Gaps 1;

QY 2 DALNVLMMNIIISKEKEIKWI 23
Db 50 DMANLFQAMNAL----KQLAWI 67

RESULT 7
BOFA BACSU STANDARD; PRT; 87 AA.
ID _BOFA BACSU STANDARD; PRT; 87 AA.
AC P24282;
DT 01-MAR-1992 (Rel. 21, Created)
DT 01-MAR-1992 (Rel. 21, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Sigma-K factor processing regulatory protein BOFA (Bypass-of-forespore
DE protein A).
GN BOFA.
OS Bacillus subtilis.
OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.
OX NCBI_TaxID=1423;
RN [1]
RN SEQUENCE FROM N.A.
RX MEDLINE=91088245; PubMed=2124672;
RA Alonso C., Shirahige K., Ogasawara N.;
RT "Molecular cloning, genetic characterization and DNA sequence
RT analysis of the recM region of Bacillus subtilis.";
RL Nucleic Acids Res. 18:6771-6777(1990).
RN [2]
RN SEQUENCE FROM N.A.
RX STRAIN=168;
RC MEDLINE=96051385; PubMed=7584024;
RA Ogasawara N., Nakai S., Yoshikawa H.;
RT "Systematic sequencing of the 180 kilobase region of the Bacillus
RT subtilis chromosome containing the replication origin.";
RL DNA Res. 1:1-14(1994).
RN [3]
RN SEQUENCE FROM N.A.
RX STRAIN=168;
RC MEDLINE=98044033; PubMed=9384377;
RA Kunst F., Ogasawara N., Moszer I., Albertini A.M., Alloni G.,

```

RA Azevedo V., Bertero M.G., Bessieres P., Bolotin A., Borchert S.,  
 RA Borries R., Boursier L., Brans A., Braun M., Brignell S.C., Bron S.,  
 RA Brouillet S., Bruschi C.V., Caldwell B., Capuano V.,  
 RA Choi S.K., Codani J.J., Connerth I.F., Cummings N.J., Daniel R.A.,  
 RA Denizot F., Devine K.M., Dusterhoft A., Ehrlich S.D., Emerson P.T.,  
 RA Entian K.D., Errington J., Fabret C., Ferrari E., Foulger D.,  
 RA Fritz C., Fujita M., Fujita Y., Fuma S., Galizzi A., Galleron N.,  
 RA Ghim S.Y., Glaeser P., Goffeau A., Gollightly E.J., Grandi G.,  
 RA Guiseppe G., Guy B.J., Haga K., Halech J., Harwood C.R., Henaut A.,  
 RA Hilbert H., Holsappel S., Honono S., Hulio M.F., Itaya M., Jones L.,  
 RA Joris B., Karamata D., Kaahara Y., Klaerr-Blanchard M., Klein C.,  
 RA Kobayashi Y., Koetter P., Koningsstein G., Krogh S., Kumano M.,  
 RA Kurita K., Lapidis A., Lardinois S., Lauber J., Lazarevic V.,  
 RA Lee S.M., Levine A., Liu H., Maunda S., Mauel C., Medigue C.,  
 RA Medina N., Mellado R.P., Mizuno M., Moestl D., Nakai S., Noback M.,  
 RA Neone D., O'Reilly M., Ogawa K., Ogiwara A., Oudega B., Park S.H.,  
 RA Parro V., Pohl T.M., Portetelle D., Rapoport G., Prescott A.M.,  
 RA Presecan E., Pujić P., Purnelle B., Rapoport G., Rey M., Reynolds S.,  
 RA Rieger M., Rivolta C., Rocha E., Roche B., Rose M., Sadaie Y.,  
 RA Sato T., Scanlan E., Schleich S., Schroeter R., Scoffone F.,  
 RA Sekiguchi J., Sekowska A., Seror S.J., Serror P., Shin B.S., Soldo B.,  
 RA Sorokin A., Tacconi E., Takagi T., Takahashi H., Takemaru K.,  
 RA Takeuchi M., Tanakoshi A., Tanaka T., Terpetra P., Tognoni A.,  
 RA Tosato V., Uchiyama S., Vandenbol M., Vannier F., Vassarotti A.,  
 RA Viari A., Wamburt R., Wedler E., Wedler H., Weitzenegger T.,  
 RA Winters P., Wiput A., Yamamoto H., Yamane K., Yasumoto K., Yata K.,  
 RA Yoshida K., Yoshikawa H.F., Zmstein E., Yoshikawa H., Danchin A.,  
 RA "The complete genome sequence of the Gram-positive bacterium Bacillus  
 RT subtilis.";  
 RL Nature 390:249-256(1997).  
 RN [4]  
 RP CHARACTERIZATION  
 RX MEDLINE=92250411; PubMed=1577688;  
 RA Ricca E., Cutting S.M., Losick R.;  
 RT "Characterization of bofA, a gene involved in intercompartmental  
 RT regulation of pro-sigma K processing during sporulation in Bacillus  
 RT subtilis.";  
 RL J. Bacteriol. 174:3177-3184(1992).  
 RN [5]  
 RP TOPOLOGY.  
 RX MEDLINE=97286525; PubMed=9141672;  
 RA Varamonti M., Marasco R., de Felice M., Sacco M.;  
 RT "Membrane topology analysis of the Bacillus subtilis BofA protein  
 RT involved in pro-sigma K processing.";  
 RL Microbiology 143:1053-1058(1997).  
 CC -1- FUNCTION: INVOLVED IN THE MEDIATION OF THE INTERCOMPARTMENTAL  
 CC COUPLING OF PRO-SIGMA K PROCESSING TO EVENTS IN THE FORESPORE.  
 CC -1- SUBCELLULAR LOCATION: Integral membrane protein.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL; X17014; CAA34880.1; -;  
 DR EMBL; D26185; BAA05255.1; -;  
 DR EMBL; Z99104; CAB11799.1; -;  
 DR PIR; B41869; B41869.  
 DR Subtilist; BG10087; bofa.  
 KW Sporulation; Transmembrane; Complete proteome.  
 FT DOMAIN 1 2 EXTRACELLULAR (PROBABLE).  
 FT TRANSMEM 3 23 PROBABLE.  
 FT DOMAIN 24 36 CYTOPLASMIC (PROBABLE).  
 FT TRANSMEM 37 57 PROBABLE.  
 FT DOMAIN 58 87 EXTRACELLULAR (PROBABLE).  
 SQ SEQUENCE 87 AA; 9010 MW; FF1DA14B5626B70E CRC64;  
 Query Match 24.3%; Score 37; DB 1; Length 87;  
 Best Local Similarity 27.3%; Pred. No. 1.2e+02;  
 Matches 6; Conservative 7; Mismatches 9; Indels 0; Gaps 0;

QY 4 LNLVLMANNIISKEKEIKWIGL 25  
 |:::|:|:|:|:  
 Db 11 LGLVILLFLSGSAAPKLKWI 32  
 |:::|:|:|:|:  
 RESULT 8  
 RL29 SULTO  
 ID RL29 SULTO STANDARD; PRT; 88 AA.  
 AC Q975f8;  
 DT 28-FEB-2003 (Rel. 41, Created)  
 DT 28-FEB-2003 (Rel. 41, Last sequence update)  
 DT 28-FEB-2003 (Rel. 41, Last annotation update)  
 DE 50S ribosomal protein L29P.  
 GN RPL29P OR STS061.  
 OS Sulfolobus tokodaii.  
 OC Archaea; Crenarchaeota; Thermoprotei; Sulfolobales; Sulfolobaceae;  
 OC Sulfolobus.  
 OC NCBI\_TaxID=111955;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=JCM 10545 / 7;  
 RX MEDLINE=21456156; PubMed=11572479;  
 RA Kawarayashi Y., Hino Y., Horikawa H., Jin-no K., Takahashi M.,  
 RA Sekine M., Baba S.-I., Ankaï A., Kosugi H., Hosoyana A., Fukui S.,  
 RA Nagai Y., Nishijima K., Otsuka R., Nakazawa H., Takamiya M., Kato Y.,  
 RA Yoshizawa T., Tanaka T., Kudoh Y., Yamazaki J., Kushida N., Oguchi A.,  
 RA Aoki K.-I., Masuda S., Yanagii M., Nishimura M., Yamagishi A.,  
 RA Oshima T., Kikuchi H.;  
 RT "Complete genome sequence of an aerobic thermoacidophilic  
 RT Crenarchaeon, Sulfolobus tokodaii strain 7.";  
 RL DNA Res. 8:123-140(2001).  
 CC -1- SIMILARITY: BELONGS TO THE L29P FAMILY OF RIBOSOMAL PROTEINS.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL; AP000982; BAB65412.1; -;  
 DR HAMAP; MF\_00374; atypical; 1;  
 DR InterPro; IPR001854; Ribosomal\_L29.  
 DR Pfam; PF00831; Ribosomal\_L29; 1.  
 DR TIGRFAMs; TIGR00012; L29; 2.  
 DR PROSITE; PS00579; RIBOSOMAL\_L29; 1.  
 DR Ribosomal protein; Complete proteome.  
 KW Ribosomal protein; Complete proteome.  
 SQ SEQUENCE 88 AA; 10277 MW; 9A5B120EB10382C2 CRC64;  
 Query Match 24.3%; Score 37; DB 1; Length 88;  
 Best Local Similarity 41.7%; Pred. No. 1.3e+02;  
 Matches 10; Conservative 4; Mismatches 6; Indels 4; Gaps 1;  
 QY 2 DALNLVLMANNIISKE-----KKEIK 21  
 |:::|:|:|:|:  
 Db 64 DIARILTLIINKENKKESK 87  
 |:::|:|:|:|:  
 RESULT 9  
 HDL\_STAAM  
 ID HDL\_STAAM STANDARD; PRT; 45 AA.  
 AC P01506;  
 DT 21-JUL-1986 (Rel. 01, Created)  
 DT 01-MAY-1991 (Rel. 18, Last sequence update)  
 DT 28-FEB-2003 (Rel. 41, Last annotation update)  
 DE Delta-henolysin precursor (Delta-toxin).  
 GN HDL OR SAV2035 OR SAS065.  
 OS Staphylococcus aureus (strain M50 / ATCC 700699),  
 OS Staphylococcus aureus (strain N315), and  
 OS Staphylococcus aureus.  
 OC Bacteria; Firmicutes; Bacillales; Staphylococcus.



OX NCBI\_TaxID=158878, 158879, 1280;  
[1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=MJ50 / ATCC 700699, and N315;  
RX MEDLINE=21311952; PubMed=11418146;  
RA Kuroda M., Ohta T., Uchiyama I., Baba T., Yuzawa H., Kobayashi I.,  
RA Cui L., Oguchi A., Aoki K.-I., Nagai Y., Lian J.-Q., Ito T.,  
RA Kananori M., Matsumaru H., Maruyama A., Murakami H., Hoshoyama A.,  
RA Mizutani-Uji Y., Takahashi N.K., Sawano T., Inoue R.-I., Kaito C.,  
RA Sekimizu K., Hirakawa H., Kihara S., Goto S., Yabuzaki J.,  
RA Kanehisa M., Yamashita A., Oshima K., Furuya K., Yoshino C., Shiba T.,  
RA Hattori M., Ogasawara N., Hayashi K., Hiramatsu K.,  
RT "whole genome sequencing of methicillin-resistant Staphylococcus  
RT aureus";  
RL Lancet 357:1225-1240(2001).  
[2]  
RP SEQUENCE FROM N.A.  
RC STRAIN=NCTC 8325-4;  
RX MEDLINE=90158509; PubMed=2622452;  
RA Janson L., Loetdahl S., Arvidson S.;  
RT "identification and nucleotide sequence of the delta-lysin gene, hld,  
RT adjacent to the accessory gene regulator (agr) of Staphylococcus  
RT aureus";  
RL Mol. Gen. Genet. 219:480-485(1989).  
[3]  
RP SEQUENCE FROM N.A.  
RC STRAIN=Isolate GAL;  
RX MEDLINE=96004766; PubMed=7565609;  
RA Novick R.P., Projan S.J., Kornblum J., Ross H.F., Ji G.,  
RA Kreiwirth B., Vandenesch F., Moghazeh S.;  
RT "The agr P2 operon: an autocatalytic sensory transduction system in  
RT Staphylococcus aureus";  
RL Mol. Gen. Genet. 248:446-458(1995).  
[4]  
RP SEQUENCE FROM N.A.  
RC STRAIN=RM4220;  
RA Bischoff M.;  
RT "DNA for hld, agrB, and agrD genes of Staphylococcus aureus strain  
RT RM4220";  
RL Submitted (FEB-2000) to the EMBL/GenBank/DBJ databases.  
[5]  
RP SEQUENCE OF 20-45.  
RC STRAIN=186X;  
RX MEDLINE=80246463; PubMed=7398877;  
RA Fitton J.E., Dell A., Shaw W.V.;  
RT "The amino acid sequence of the delta haemolysin of Staphylococcus  
RT aureus";  
RL FEBS Lett. 115:209-212(1980).  
[6]  
RP SEQUENCE OF 20-45.  
RC STRAIN=Canine variant;  
RA Fitton J.E., Hunt D.F., Marasco J., Shabanowitz J., Winston S.,  
RA Dell A.;  
RT "The amino acid sequence of delta haemolysin purified from a canine  
RT isolate of S. aureus";  
RL FEBS Lett. 169:25-29(1984).  
[7]  
RP 3D-STRUCTURE MODELING.  
RX MEDLINE=91126037; PubMed=2281085;  
RA Reghunathan G., Seetharamulu P., Brooks B.R., Guy H.R.;  
RT "Models of delta-hemolysin membrane channels and crystal structures";  
RL Proteins 8:213-225(1990).  
CC -1- FUNCTION: DELTA-TOXIN Lyses ERYTHROCYTES AND MANY OTHER MAMMALIAN  
CC CELLS.  
CC -1- SUBCELLULAR LOCATION: ASSUMED TO INSERT AND OLIGOMERIZE IN  
CC MEMBRANES.  
CC  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>)  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch).  
CC

CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch).  
CC  
DR EMBL; AP003364; BAB58197.1; ALT INIT.  
DR EMBL; AP003135; BAB43122.1; ALT INIT.  
DR EMBL; X52543; CAA36780.1; ALT INIT.  
DR EMBL; AF230358; AAF43204.1; -  
DR PIR; A01767; LSSAD.  
DR PIR; A89995; A89995.  
DR PDB; 1DHL; 15-OCT-94.  
DR PDB; 2DTB; 31-OCT-93.  
DR PDB; 1DTC; 31-OCT-93.  
DR PDB; PFO5372; Delta\_lysin; 1.  
KW Hemolysis; Formylation; Toxin; Transmembrane; 3D-structure;  
KW Complete proteome.  
FT PROPEP 1 19  
FT CHAIN 20 45 DELTA-HEMOLYSIN.  
FT MOD\_RES 20 20 FORMYLATION (ABOUT 90% OF THE MOLECULES).  
FT VARIANT 22 22 Q -> A (IN CANINE STRAIN).  
FT VARIANT 29 31 GDL -> VEF (IN CANINE STRAIN).  
FT VARIANT 34 34 W -> L (IN CANINE STRAIN).  
FT VARIANT 36 37 ID -> AE (IN CANINE STRAIN).  
FT VARIANT 40 40 N -> E (IN CANINE STRAIN).  
FT VARIANT 43 43 T -> I (IN CANINE STRAIN).  
FT HELIX 21 41  
SQ SEQUENCE 45 AA; 5140 MW; 3686661E0342CA01 CRC64;  
Query Match 23.7%; Score 36; DB 1; Length 45;  
Best Local Similarity 44.4%; Pred. No. 89;  
Matches 8; Conservative 3; Mismatches 7; Indels 0; Gaps 0;  
Qy 6 VLMAMIIISKEKEIKWI 23  
Db 18 ISMAQDIISTIGDLVKWI 35  
RESULT 10  
Y546 METJA STANDARD; PRT; 56 AA.  
AC Q57966;  
DT 01-NOV-1997 (Rel. 35, Created)  
DT 01-NOV-1997 (Rel. 35, Last sequence update)  
DT 16-OCT-2001 (Rel. 40, Last annotation update)  
DE Hypothetical protein MJ0546.  
GN MJ0546.  
OS Methanococcus jannaschii.  
OC Archaea; Euryarchaeota; Methanococci; Methanococcales;  
OC Methanocaldococcales; Methanocaldococcus.  
OX NCBI\_TaxID=2190;  
RX [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=JAL-1 / DSM 2661 / ATCC 43067;  
RX MEDLINE=96337999; PubMed=8688087;  
RA Sult C.J., White O., Olsen G.J., Zhou L., Fleischmann R.D.,  
RA Sutton G.G., Blake J.A., Fitzgerald L.M., Clayton R.A., Gocayne J.D.,  
RA Krolavage A.R., Dougherty B.A., Tomb J.-F., Adams M.D., Reich C.I.,  
RA Overbeek R., Kirkness E.F., Weinstock K.G., Merrick J.M., Glodek A.,  
RA Scott J.L., Geoghegan N.S.M., Weidman J.F., Fuhrmann J.L., Nguyen D.,  
RA Uterback T.R., Kelley J.M., Peterson J.D., Sadow P.W., Hanna M.C.,  
RA Cotton M.D., Roberts K.M., Hurst M.A., Kaine B.P., Borodovsky M.,  
RA Klenk H.-P., Fraser C.M., Smith H.O., Woese C.R., Venter J.C.;  
RT "Complete genome sequence of the methanogenic archaeon, Methanococcus  
RT jannaschii";  
RL Science 273:1058-1073(1996).  
CC  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>)  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch).  
CC  
DR EMBL; U67504; AAB98548.1; -

Tue Feb 17 11:55:55 2004

```

DR PIR; B64368; B64368.
DR TIGR; M0546; -.
KW Hypothetical protein; Transmembrane; Complete proteome.
FT TRANSMEM 22 POTENTIAL.
FT DOMAIN 27 47
FT GLU-RICH 2
SQ SEQUENCE 56 AA; 6460 MW; 66CBFC395548BASD CRC64;

Query Match      23.7%; Score 36; DB 1; Length 56;
Best Local Similarity 35.0%; Pred. No. 1.1e+02;
Matches 7; Conservative 7; Mismatches 4; Indels 1; Gaps 1;

QY 4 LNVLMAMNII--SKEKKEIK 22
Db 13 LNIILGKVIQLQLEEVK 32

RESULT 11
Y295_ARCFU
ID Y295_ARCFU STANDARD; PRT; 74 AA.
AC 02947;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Hypothetical protein AF0295.
GN AF0295.
OS Archaeoglobus fulgidus.
OC Archaea; Euryarchaeota; Archaeoglobi; Archaeoglobales;
OC Archaeoglobaceae; Archaeoglobus.
OX NCBI_TaxID=2234;
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN=VC-16 / DSM 4304 / ATCC 49558;
RX MEDLINE=98049343; PubMed=9389475;
RA Klenk H.-P., Clayton R.A., Tomb J.-F., Hickey E.K., Peterson J.D.,
RA Ketchum K.A., Dodson R.J., Gwinn M., Hickey E.K., Kyrpides N.C.,
RA Richardson D.L., Kerlavage A.R., Graham D.E., Sutton G.G., Gill S.,
RA Fleischmann R.D., Quackenbush J., Lee N.H., Sutton G.G., Gill S.,
RA Kirkness E.F., Dougherty B.A., McKenney K., Adams M.D., Loftus B.,
RA Peterson S., Reich C.I., McNeil L.K., Badger J.H., Glodek A., Zhou L.,
RA Overbeek R., Gocayne J.D., Weidman J.F., McDonald L., Utterback T.,
RA Cotton M.D., Spriggs T., Artach P., Kaine B.P., Sykes S.M.,
RA Sadow P.W., D'Andrea K.P., Bowman C., Fujii C., Garland S.A.,
RA Mason T.M., Olsen G.J., Fraser C.M., Smith H.O., Woese C.R.,
RA Venter J.C.;
RA "The complete genome sequence of the hyperthermophilic, sulphate-
RT reducing archaeon Archaeoglobus fulgidus."
RL Nature 390:364-370 (1997).
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; AB001084; AAB90939.1; -.
DR PIR; G69286; G69286.
DR TIGR; AF0295; -.
KW Hypothetical protein; Complete proteome.
SQ SEQUENCE 74 AA; 8839 MW; 0DCF5EF3B444B2F5 CRC64;

Query Match      23.7%; Score 36; DB 1; Length 74;
Best Local Similarity 47.4%; Pred. No. 1.5e+02;
Matches 9; Conservative 2; Mismatches 8; Indels 0; Gaps 0;

QY 4 LNVLMAMNIIISKEKKEIKW 22
Db 3 VQVLEAGKIISPNEKVIW 21

RESULT 12
YDFK_ECOLI
ID YNAE_ECOLI STANDARD; PRT; 88 AA.
AC P76154;
DT 15-JUL-1998 (Rel. 36, Created)
DT 15-JUL-1998 (Rel. 36, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE Hypothetical protein ydfk.
GN YDFK OR B1544.
OS Escherichia coli.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Escherichia.
OX NCBI_TaxID=562;
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN=K12 / MG1655;
RX MEDLINE=97426617; PubMed=9278503;
RA Blattner F.R., Plunkett G. III, Bloch C.A., Perna N.T., Burland V.,
RA Riley M., Collado-Vides J., Glasner J.D., Rode C.K., Mayhew G.F.,
RA Gregor J., Davis N.W., Kirkpatrick H.A., Goeden M.A., Rose D.J.,
RA Mau B., Shao Y.;
RA "The complete genome sequence of Escherichia coli K-12."
RL Science 277:1453-1474 (1997).
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC -----
CC EMBL; AF000252; AAC74617.1; -.
DR PIR; C64909; C64909.
DR EcoGene; EG13823; ydfk.
KW Hypothetical protein; Complete proteome.
SQ SEQUENCE 88 AA; 10136 MW; 534B06E2287CC6A0 CRC64;

Query Match      23.7%; Score 36; DB 1; Length 88;
Best Local Similarity 44.4%; Pred. No. 1.8e+02;
Matches 8; Conservative 2; Mismatches 6; Indels 2; Gaps 1;

QY 5 NVLMAMNIIISKEKKEIKW 22
Db 5 NIL--RNIFMKSKDTLKW 20

RESULT 13
YNAE_ECOLI
ID YNAE_ECOLI STANDARD; PRT; 88 AA.
AC P76073;
DT 15-JUL-1998 (Rel. 36, Created)
DT 15-JUL-1998 (Rel. 36, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE Hypothetical protein ynae.
GN YNAE OR B1375.
OS Escherichia coli.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Escherichia.
OX NCBI_TaxID=562;
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN=K12 / MG1655;
RX MEDLINE=97426617; PubMed=9278503;
RA Blattner F.R., Plunkett G. III, Bloch C.A., Perna N.T., Burland V.,
RA Riley M., Collado-Vides J., Glasner J.D., Rode C.K., Mayhew G.F.,
RA Gregor J., Davis N.W., Kirkpatrick H.A., Goeden M.A., Rose D.J.,
RA Mau B., Shao Y.;
RA "The complete genome sequence of Escherichia coli K-12."
RL Science 277:1453-1474 (1997).
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC -----

```

CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

CC -----  
 DR EMBL; A8000234; AAC74457.1; -  
 DR PIR; B64888; B64888.  
 DR EcoGene; EGI3373; ynaB.  
 KW Hypothetical protein; Complete proteome.  
 SQ SEQUENCE 88 AA; 10109 MW; 534B1813C77CC6A0 CRC64;

Query Match 23.7%; Score 36; DB 1; Length 88;  
 Best Local Similarity 44.4%; Pred. No. 1.9e+02;  
 Matches 8; Conservative 2; Mismatches 6; Indels 2; Gaps 1;

QY 5 NVLMAMNIISKEKKEIKW 22  
 :|: ||| :|: |||  
 Db 5 NII--RNIFPKSKDTLKW 20

## RESULT 14

YOEI BACSU STANDARD; PRT; 96 AA.  
 AC F54454;

DT 01-OCT-1996 (Rel. 34, Created)  
 DT 01-OCT-1996 (Rel. 34, Last sequence update)  
 DT 28-FEB-2003 (Rel. 41, Last annotation update)  
 DE Hypothetical protein yqeI.  
 GN YOEI.

OS Bacillus subtilis.

OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.

OX NCBI\_TaxID=1423;

[1]

RP SEQUENCE FROM N.A.

RC STRAIN=168 / JH642;

RX MEDLINE=97124195; PubMed=8969508;

RA Mizuno M., Masuda S., Takemaru K.-I., Hosono S., Sato T., Takeuchi M.,  
 RA Kobayashi Y.;

RT "Systematic sequencing of the 283 kb 210 degrees-232 degrees region of  
 RT the Bacillus subtilis genome containing the skin element and many  
 RT sporulation genes.";  
 RL Microbiology 142:3103-3111 (1996).

[2]

RP SEQUENCE FROM N.A.

RC STRAIN=168;

RX MEDLINE=98044033; PubMed=9384377;

RA Kunst F., Ogasawara N., Moszer I., Albertini A.M., Alloni G.,  
 RA Azevedo V., Bertero M.G., Bessieres P., Bolotin A., Borchert S.,  
 RA Borriess R., Boursier L., Brans A., Braun M., Brignell S.C., Bron S.,  
 RA Brouillet S., Bruschi C.V., Caldwell B., Capuano V., Carter N.M.,  
 RA Choi S.K., Codani J.J., Connerton I.F., Cummings N.J., Daniel R.A.,  
 RA Denizot P., Devine K.M., Dusterhoft A., Ehrlich S.D., Emerson P.T.,  
 RA Entian K.D., Errington J., Fabre C., Ferrari E., Foulger D.,  
 RA Fritz C., Fujita M., Fujita Y., Fuma S., Gallizzi A., Galleron N.,  
 RA Chim S.Y., Glaser P., Goffeau A., Gollightly E.J., Grandi G.,  
 RA Guisepi G., Guy B.J., Haga K., Haiech J., Harwood C.R., Henaut A.,  
 RA Hilbert H., Holsappel S., Hosono S., Hullo M.F., Itava M., Jones L.,  
 RA Joris B., Karamata D., Kasahara Y., Klaerr-Blanchard M., Klein C.,  
 RA Kobayashi Y., Koetter P., Koningsstein G., Krogh S., Kumano M.,  
 RA Kurita K., Lapidus A., Lardinois S., Lauber J., Lazarevic V.,  
 RA Lee S.M., Levine A., Liu H., Masuda S., Maue C., Medigue C.,  
 RA Medina N., Mellado R.P., Mizuno M., Moestl D., Nakai S., Noback M.,  
 RA Noone D., O'Reilly M., Ogawa K., Ogiwara A., Oudega B., Park S.H.,  
 RA Parro V., Pohl T.M., Portetelle D., Porwollik S., Prescott A.M.,  
 RA Presecan E., Fujic P., Purnelle B., Rapoport G., Rey M., Reynolds S.,  
 RA Rieger M., Rivolta C., Rocha E., Roche B., Rose M., Sadale Y.,  
 RA Sato T., Scanlan E., Schleich S., Schroeter R., Scoffone F.,  
 RA Sekiguchi J., Sekowska A., Seror S.J., Serror P., Shin B.S., Soldo B.,  
 RA Sorokin A., Tacconi E., Takagi T., Takahashi H., Takemaru K.,  
 RA Takeuchi M., Tamakoshi A., Tanaka T., Terpetra P., Tognoni A.,  
 RA Tosato V., Uchiyama S., Vandenbol M., Vannier F., Vassarotti A.,  
 RA Viari A., Wambutt R., Wedler E., Wedler H., Weitzenecker T.,

RA Winters P., Wipat A., Yamamoto H., Yamane K., Yasumoto K., Yata K.,  
 RA Yoshida K., Yoshikawa H.F., Zumstein E., Yoshikawa H., Danchin A.,  
 RT "The complete genome sequence of the Gram-positive bacterium Bacillus  
 RT subtilis.";  
 RL Nature 390:249-256 (1997).

CC -1- SIMILARITY: BELONGS TO THE UPF0044 FAMILY.

CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

CC -----  
 DR EMBL; D84432; BAA12446.1; -  
 DR EMBL; Z99117; CAB14507.1; -  
 DR PIR; E69951; E69951.  
 DR Subtilist; EGI1637; yqeI.  
 DR InterPro; IPR001890; UPF0044.  
 DR Pfam; PF01985; UPF0044; 1.  
 DR ProDom; PD010559; UPF0044; 1.  
 DR TIGRfam; TIGR00253; TIGR00253; 1.  
 DR PROSITE; PS01301; UPF0044; 1.  
 KW Hypothetical protein; Complete proteome.  
 SQ SEQUENCE 96 AA; 10758 MW; D9BA3E935C161F54 CRC64;

Query Match 23.7%; Score 36; DB 1; Length 96;

Best Local Similarity 53.3%; Pred. No. 1.9e+02;  
 Matches 8; Conservative 2; Mismatches 5; Indels 0; Gaps 0;

QY 12 IISKEKKEIKWIGLP 26  
 :|: ||| :|: |||  
 Db 82 VLYKESKKNQIPLP 96

## RESULT 15

YHBY ECOLI

ID YHBY ECOLI STANDARD; PRT; 97 AA.

AC P42550;

DT 01-NOV-1995 (Rel. 32, Created)

DT 01-NOV-1995 (Rel. 32, Last sequence update)

DT 28-FEB-2003 (Rel. 41, Last annotation update)

DE Hypothetical protein yhbY.

GN YHBY OR B3180 OR C3937 OR Z4542 OR ECS4059 OR SF3220.

OS Escherichia coli.

OS Escherichia coli O6

OS Escherichia coli O157:H7, and

OS Shigella flexneri.

OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;

OC Enterobacteriaceae; Escherichia.

OX NCBI\_TaxID=562, 217992, 83334, 623;

RN [1]

RP SEQUENCE FROM N.A.

RC SPECIES=E.coli; STRAIN=K12 / MG1655;

RX MEDLINE=97426617; PubMed=9278503;

RA Blattner F.R., Plunkett G. III, Bloch C.A., Perna N.T., Burland V.,

RA Riley M., Collado-Vides J., Glaesner J.D., Rode C.K., Mayhew G.P.,

RA Gregor J., Davis N.W., Kirkpatrick H.A., Goeden M.A., Rose D.J.,

RA Mau B., Shao Y.;

RT "The complete genome sequence of Escherichia coli K-12.";

RL Science 277:1453-1474 (1997).

[2]

RN SEQUENCE FROM N.A.

RC SPECIES=E.coli; STRAIN=K12;

RA Wang R., Kushner S.R.;

RL Submitted (SEP-1993) to the EMBL/GenBank/DBJ databases.

[3]

RN SEQUENCE FROM N.A.

RC SPECIES=E.coli; STRAIN=O6:H1 / CFT073 / ATCC 700928;

RX MEDLINE=22388224; PubMed=12471157;

RA Welch R.A., Burland V., Plunkett G. III, Redford P., Rosch P.,

RA Rasko D., Buckles E.L., Liou S.-R., Boutin A., Hackett J., Stroud D.,

DR Pfam; PF01985; UPF0044; 1.  
DR ProDom; PD010559; UPF0044; 1.  
DR TIGRFAMs; TIGR00253; TIGR00253; 1.  
DR PROSITE; PS01301; UPF0044; 1.  
KW Hypothetical protein; Complete proteome.  
SQ SEQUENCE 97 AA; 10784 MW; 52B4F7AD4C203382 CRC64;

Query Match 23.7%; Score 36; DB 1; Length 97;  
Best Local Similarity 35.3%; Pred. No. 2e+02;  
Matches 6; Conservative 6; Mismatches 5; Indels 0; Gaps 0;

OY 10 MNIISKEKIKWIGLP 26  
||:|:|:|:|:|:|  
DB 1 MNLSTKQKHLKGLAHP 17

RESULT 16  
TRPC\_BACPU STANDARD; PRT; 68 AA.  
ID TRPC\_BACPU STANDARD; PRT; 68 AA.  
AC P18268;  
DT 01-NOV-1990 (Rel. 16, Created)  
DT 01-NOV-1990 (Rel. 16, Last sequence update)  
DT 28-FEB-2003 (Rel. 41, Last annotation update)  
DE Indole-3-glycerol phosphate synthase (EC 4.1.1.48) (IGPS) (Fragment).  
GN TRPC.  
OS Bacillus pumilus (Bacillus mesentericus).  
OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.  
OX NCBI\_TaxID=1408;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=RU502;  
RX MEDLINE=90236301; PubMed=2110100;  
RA Rivas M.V., Jarvis E.D., Rudner R.;  
RT "The structure of the trpE, trpD and 5' trpC genes of Bacillus  
RT pumilus."  
RL Gene 87:71-78 (1990).  
RN [2]  
RP ERRATUM.  
RX MEDLINE=91033058; PubMed=2227447;  
RA Rivas M.V., Jarvis E.D., Rudner R.;  
RL Gene 94:141-143 (1990).  
CC -! CATALYTIC ACTIVITY: 1-(2-carboxyphenylamino)-1-deoxy-D-ribulose 5-  
CC phosphate = 1-(indol-3-yl)glycerol 3-phosphate + CO(2) + H(2)O.  
CC -! PATHWAY: Tryptophan biosynthesis; fourth step.  
CC -! SIMILARITY: BELONGS TO THE TRPC FAMILY.  
CC  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC  
CC EMBL; M36468; AAB02274.1; -  
CC PIR; JH0100; JH0100.  
DR HAMAP; MF 00134; -; 1.  
DR InterPro; IPR001468; IGPS.  
DR Pfam; PF00218; IGPS; 1.  
DR ProDom; PD001511; IGPS; 1.  
DR PROSITE; PS00614; IGPS; 1.  
KW Tryptophan biosynthesis; Lyase; Decarboxylase.  
FT NON TER 68  
SQ SEQUENCE 68 AA; 7688 MW; DFDDBEFC80E7BDE4 CRC64;

Query Match 23.0%; Score 35; DB 1; Length 68;  
Best Local Similarity 33.3%; Pred. No. 1.9e+02;  
Matches 7; Conservative 7; Mismatches 7; Indels 0; Gaps 0;

OY 8 MAMNIISKEKIKWIGLPTN 28  
||:|:|:|:|:|:|  
DB 1 MLNQIIRKKEHIGTLQLPVD 21

Mayhew G.F., Rose D.J., Zhou S., Schwartz D.C., Perna N.T.,  
Mobley H.L.T., Sonnenberg M.S., Blattner F.R.;  
"Extensive mosaic structure revealed by the complete genome sequence  
of uropathogenic Escherichia coli";  
Proc. Natl. Acad. Sci. U.S.A. 99:17020-17024 (2002).  
[4]  
SEQUENCE FROM N.A.  
SPECIES=E.coli; STRAIN=O157:H7 / EDL933 / ATCC 700927;  
MEDLINE=21074935; PubMed=11206551;  
Perna N.T., Plunkett G. III, Burland V., Mau B., Glasner J.D.,  
Rose D.J., Mayhew G.F., Evans P.S., Gregor J., Kirkpatrick H.A.,  
Posfal G., Hackett J., Klink S., Boutin A., Shao Y., Miller L.,  
Grotbeck E.J., Davis N.W., Lim A., Dimlanta E.T., Potamouis K.,  
Apodaca J., Anantharaman T.S., Lin J., Yen G., Schwartz D.C.,  
Welch R.A., Blattner F.R.;  
"Genome sequence of enterohaemorrhagic Escherichia coli O157:H7";  
Nature 409:529-533 (2001).  
[5]  
SEQUENCE FROM N.A.  
SPECIES=E.coli; STRAIN=O157:H7 / RIMD 0509952;  
MEDLINE=21156231; PubMed=11258796;  
Hayashi T., Makino K., Ohnishi M., Kurokawa K., Ishii K., Yokoyama K.,  
Han C.-G., Ohtsubo E., Nakayama K., Murata T., Tanaka M., Tobe T.,  
Iida T., Takami H., Honda T., Saeakawa C., Ogasawara N., Yasunaga T.,  
Kuhara S., Shiba T., Hattori M., Shinagawa H.;  
"Complete genome sequence of enterohaemorrhagic Escherichia coli  
O157:H7 and genomic comparison with a laboratory strain K-12";  
DNA Res. 8:11-22 (2001).  
[6]  
SEQUENCE OF 64-97 FROM N.A.  
SPECIES=E.coli;  
MEDLINE=91057148; PubMed=2243801;  
Sparks J., Das A.;  
RT "The nucleotide sequence of greA, a suppressor gene that restores  
RT growth of an Escherichia coli RNA polymerase mutant at high  
RT temperature";  
RL Nucleic Acids Res. 18:6443-6443 (1990).  
[7]  
SEQUENCE FROM N.A.  
SPECIES=flexneri; STRAIN=301 / Serotype 2a;  
MEDLINE=22272406; PubMed=13384590;  
Jin Q., Yuan Z., Xu J., Wang Y., Shen Y., Lu W., Wang J., Liu H.,  
Yang J., Yang F., Zhang X., Zhang J., Yang G., Wu H., Qu D., Dong J.,  
Sun L., Xue Y., Zhao A., Gao Y., Zhu J., Kan B., Ding K., Chen S.,  
Cheng H., Yao Z., He B., Chen R., Ma D., Qiang B., Hou Y.,  
Yu J.;  
"Genome sequence of Shigella flexneri 2a: insights into pathogenicity  
through comparison with genomes of Escherichia coli K12 and O157";  
Nucleic Acids Res. 30:4432-4444 (2002).  
[8]  
-! SIMILARITY: BELONGS TO THE UPF0044 FAMILY. STRONG, TO H.INFLUENZA  
[9]  
H11333.  
CC  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC  
CC EMBL; U18997; AAS7981.1; -  
CC EMBL; AE000398; AAC76212.1; -  
CC EMBL; U01376; -; NOT ANNOTATED\_CDS.  
CC EMBL; AE016767; AAN82377.1; -  
CC EMBL; AE005546; AAG58314.1; -  
CC EMBL; AP002564; BAB37482.1; -  
CC EMBL; X54718; -; NOT ANNOTATED\_CDS.  
CC EMBL; AE015333; ABA4666.1; -  
CC PIR; C91136; C91136.  
CC PIR; F65108; F65108.  
CC PIR; F85981; F85981.  
CC EcoGene; EG12794; yhbY.  
CC InterPro; IPR001890; UPF0044.  
CC

```

RESULT 17
COX2_LEPSP
ID_COX2_LEPSP STANDARD; PRT; 72 AA.
AC P29656;
DT 01-APR-1993 (Rel. 25, Created)
DT 01-APR-1993 (Rel. 25, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Cytochrome c oxidase polypeptide II (EC 1.9.3.1) (Fragment).
GN COXII OR COII.
OS Lepisosteus oculatus (Spotted gar).
OG Mitochondrion.
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Semionotiformes; Lepisosteidae;
OC Lepisosteus.
OX NCBI_TaxID=7918;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=92130804; PubMed=1663569;
RA Normark B.B., McCune A.R., Harrison R.G.;
RT "Phylogenetic relationships of neopterygian fishes, inferred from
  mitochondrial DNA sequences."
RL Mol. Biol. Evol. 8:819-834(1991).
CC -!- FUNCTION: Cytochrome c oxidase is the component of the respiratory
  chain that catalyzes the reduction of oxygen to water. Subunits 1-
  3 form the functional core of the enzyme complex. Subunit 2
  transfers the electrons from cytochrome c via its binuclear copper
  A center to the bimetallic center of the catalytic subunit 1.
CC -!- CATALYTIC ACTIVITY: 4 ferrocycytochrome c + O(2) = 4 ferricytochrome
  c + 2 H(2)O.
CC -!- COPACTOR: Copper A.
CC -!- SUBCELLULAR LOCATION: Integral membrane protein. Mitochondrial
  inner membrane.
CC -!- SIMILARITY: Belongs to the cytochrome c oxidase subunit 2 family.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
  between the Swiss Institute of Bioinformatics and the EMBL outstation -
  the European Bioinformatics Institute. There are no restrictions on its
  use by non-profit institutions as long as its content is in no way
  modified and this statement is not removed. Usage by and for commercial
  entities requires a license agreement (See http://www.isb-sib.ch/announce/
  or send an email to license@isb-sib.ch).
CC
EMBL; M64898; AAB01464.1; --
DR EMBL; M64901; AAB01467.1; --
DR InterPro; IPR001505; Copper CuA.
DR InterPro; IPR002429; Cyt_c_ox_2.
DR Pfam; PF02790; COX2_TW; 1.
DR PROSITE; PS00078; COX2; PARTIAL.
KW Oxidoreductase; Copper; Mitochondrion; Transmembrane; Inner membrane;
KW Electron transport; Respiratory chain.
FT DOMAIN 1 26 MITOCHONDRIAL INTERMEMBRANE (POTENTIAL).
FT TRANSMEM 27 48 POTENTIAL.
FT DOMAIN 49 62 MITOCHONDRIAL MATRIX (POTENTIAL).
FT TRANSMEM 63 >72 POTENTIAL.
FT NON_TER 72 72
SQ SEQUENCE 72 AA; 8081 MW; AABC3332F740FA32 CRC64;

Query Match 23.0%; Score 35; DB 1; Length 72;
Best Local Similarity 29.0%; Pred. No. 2e+02;
Matches 9; Conservative 7; Mismatches 7; Indels 8; Gaps 1;

QY 4 LNLVLMAM-----NIISKEKEIKWIGLP 26
Db 39 LYIIIVAVSTKLTNKHLDQVEIVWTILP 69

RESULT 18
COX2_LEPSP
ID_COX2_LEPSP STANDARD; PRT; 72 AA.
AC P29656;
DT 01-APR-1993 (Rel. 25, Created)
DT 01-APR-1993 (Rel. 25, Last sequence update)

```

```

DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Cytochrome c oxidase polypeptide II (EC 1.9.3.1) (Fragment).
GN COXII OR COII.
OS Lepisosteus spatula (Alligator gar) (Atractosteus spatula).
OG Mitochondrion.
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Semionotiformes; Lepisosteidae;
OC Lepisosteus.
OX NCBI_TaxID=7917;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=92130804; PubMed=1663569;
RA Normark B.B., McCune A.R., Harrison R.G.;
RT "Phylogenetic relationships of neopterygian fishes, inferred from
  mitochondrial DNA sequences."
RL Mol. Biol. Evol. 8:819-834(1991).
CC -!- FUNCTION: Cytochrome c oxidase is the component of the respiratory
  chain that catalyzes the reduction of oxygen to water. Subunits 1-
  3 form the functional core of the enzyme complex. Subunit 2
  transfers the electrons from cytochrome c via its binuclear copper
  A center to the bimetallic center of the catalytic subunit 1.
CC -!- CATALYTIC ACTIVITY: 4 ferrocycytochrome c + O(2) = 4 ferricytochrome
  c + 2 H(2)O.
CC -!- COPACTOR: Copper A.
CC -!- SUBCELLULAR LOCATION: Integral membrane protein. Mitochondrial
  inner membrane.
CC -!- SIMILARITY: Belongs to the cytochrome c oxidase subunit 2 family.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
  between the Swiss Institute of Bioinformatics and the EMBL outstation -
  the European Bioinformatics Institute. There are no restrictions on its
  use by non-profit institutions as long as its content is in no way
  modified and this statement is not removed. Usage by and for commercial
  entities requires a license agreement (See http://www.isb-sib.ch/announce/
  or send an email to license@isb-sib.ch).
CC
EMBL; M64899; AAB01455.1; ALT SEQ.
DR InterPro; IPR001505; Copper CuA.
DR InterPro; IPR002429; Cyt_c_ox_2.
DR Pfam; PF02790; COX2_TW; 1.
DR PROSITE; PS00078; COX2; PARTIAL.
KW Oxidoreductase; Copper; Mitochondrion; Transmembrane; Inner membrane;
KW Electron transport; Respiratory chain.
FT DOMAIN 1 26 MITOCHONDRIAL INTERMEMBRANE (POTENTIAL).
FT TRANSMEM 27 48 POTENTIAL.
FT DOMAIN 49 62 MITOCHONDRIAL MATRIX (POTENTIAL).
FT TRANSMEM 63 >72 POTENTIAL.
FT NON_TER 72 72
SQ SEQUENCE 72 AA; 8081 MW; AABC3332F740FA32 CRC64;

Query Match 23.0%; Score 35; DB 1; Length 72;
Best Local Similarity 29.0%; Pred. No. 2e+02;
Matches 9; Conservative 7; Mismatches 7; Indels 8; Gaps 1;

QY 4 LNLVLMAM-----NIISKEKEIKWIGLP 26
Db 39 LYIIIVAVSTKLTNKHLDQVEIVWTILP 69

RESULT 19
FCEG_MOUSE
ID_FCEG_MOUSE STANDARD; PRT; 86 AA.
AC P20491;
DT 01-FEB-1991 (Rel. 17, Created)
DT 01-FEB-1991 (Rel. 17, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE High affinity immunoglobulin epsilon receptor gamma-subunit precursor
  (FcER1) (IGE Fc receptor gamma-subunit) (Fc-epsilon RI-gamma).
GN FCE1G OR FCE1G.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;

```

```

RN  SEQUENCE FROM N.A.
RP  MEDLINE=89359361; PubMed=2527850;
RA  Ra C., Jouvin M.H.E., Kinet J.-P.;
RT  "Complete structure of the mouse mast cell receptor for IgE (Fc
RT  epsilon RI) and surface expression of chimeric receptors
RT  (rat-mouse-human) on transfected cells.";
RL  J. Biol. Chem. 264:15323-15327(1989).
CC  -1- FUNCTION: THE GAMMA SUBUNIT HAS A CRITICAL ROLE IN ALLOWING THE
CC  IGE FC RECEPTOR TO REACH THE CELL SURFACE.
CC  -1- SUBUNIT: IGE FC RECEPTOR IS A TETRAMER OF AN ALPHA CHAIN, A BETA
CC  CHAIN, AND TWO DISULFIDE LINKED GAMMA CHAINS. THE GAMMA CHAIN FROM
CC  FC-EPSILON-RI ARE ALSO SUBUNITS OF OTHER FC RECEPTORS.
CC  -1- SUBCELLULAR LOCATION: Type I membrane protein.
CC  -1- SIMILARITY: BELONGS TO THE CD32/FCER1G FAMILY.
CC  -1- SIMILARITY: CONTAINS 1 ITAM DOMAIN.
CC  This SWISS-PROT entry is copyright. It is produced through a collaboration
CC  between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC  the European Bioinformatics Institute. There are no restrictions on its
CC  use by non-profit institutions as long as its content is in no way
CC  modified and this statement is not removed. Usage by and for commercial
CC  entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC  or send an email to license@isb-sib.ch).
-----
DR  EMBL: J05020; AAA37602.1; -.
DR  MGD; MG1:95496; Fcεr1g.
DR  InterPro; IPR003110; ITAM.
DR  Pfam; PF02189; ITAM; 1.
DR  SMART; SM00077; ITAM; 1.
DR  IGE-binding protein; Receptor; Transmembrane; Signal.
FT  SIGNAL 1 18
FT  CHAIN 19 86
FT  DOMAIN 19 23
FT  TRANSMEM 24 44
FT  DOMAIN 45 86
FT  DOMAIN 62 82
FT  DISULFID 25 25
FT  SEQUENCE 86 AA; 9652 MW; 83184DE22FCC9ECB CRC64;
SQ
Query Match 23.0%; Score 35; DB 1; Length 86;
Best Local Similarity 44.4%; Pred. No. 2.5e+02;
Matches 8; Conservative 3; Mismatches 7; Indels 0; Gaps 0;

QY 12 IISKEKEIKWIGLPTNS 29
Db 55 IASREKADAVYTGILNTRS 72
-----
RESULT 20
Y511_BUCAP STANDARD; PRT; 95 AA.
AC Q8K946;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Hypothetical protein Bueg511.
GN BUEG511.
OS Buchnera aphidicola (subsp. Schizaphis graminum).
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Buchnera.
OX NCBI_TaxID=98794;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=22084549; PubMed=12089438;
RA Tanas I., Kjaersom L., Canbaeck B., Naeslund A.K., Eriksson A.-S.,
RA Wernegreen J.J., Sandstroem J.P., Moran N.A., Andersson S.G.E.;
RT "50 million years of genomic stasis in endosymbiotic bacteria.";
RL Science 296:2376-2379(2002).
CC -1- SIMILARITY: STRONG, TO E.COLI YHEL.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
-----
CC

```

```
ID NU3M_NEUCR STANDARD; PRT; 52 AA.
AC Q5141;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE NADH-ubiquinone oxidoreductase chain 3 (EC 1.6.5.3) (Fragment).
GN ND3
OS Neurospora crassa.
OG Mitochondrion.
OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Sordariomycetes;
OC Sordariomycetidae; Sordariales; Sordariaceae; Neurospora.
OX NCBI_TaxID=5141;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=86220159; PubMed=3011426;
RA de Vries H., Alzner-Dewerd B., Breitenberger C.A., Chang D.D.,
RA de Jonge J.C., Rajbhandary U.L.;
RT "The E35 stopper mutant of Neurospora crassa: precise localization of
RT deletion endpoints in mitochondrial DNA and evidence that the deleted
RT DNA codes for a subunit of NADH dehydrogenase.";
RL EMOB J. 5:779-785(1986).
CC -1- CATALYTIC ACTIVITY: NADH + ubiquinone = NAD(+) + ubiquinol.
CC -1- SIMILARITY: BELONGS TO THE COMPLEX I SUBUNIT 3 FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announcement/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; X03793; CAA27419.1; -
KW Oxidoreductase; NAD; Ubiquinone; Mitochondrion; Transmembrane.
FT NON TPR 52
FT FT 52
SQ SEQUENCE 52 AA; 5952 MW; B605E03BF203B391 CRC64;

Query Match 22.7%; Score 34.5; DB 1; Length 52;
Best Local Similarity 30.8%; Pred. No. 1.7e+02;
Matches 8; Conservative 6; Mismatches 11; Indels 1; Gaps 1;

QY 4 LNLVLMANNIISK-EKKEIKWIGLPTN 28
Dd 21 INLVFAPHIPKNWNGKSNVVKLENS 46

RESULT 23
RS16_STR33 STANDARD; PRT; 90 AA.
AC Q8E4H3; Q8DYW5;
DT 15-SEP-2003 (Rel. 42, Created)
DT 15-SEP-2003 (Rel. 42, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE 30S ribosomal protein S16.
GN RPS16 OR GBS1428 OR SAG1358.
OS Streptococcus agalactiae (serotype III), and
OS Streptococcus agalactiae (serotype V).
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC Streptococcus.
OX NCBI_TaxID=216495, 216466;
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN=NEW316 / Serotype III;
RX MEDLINE=2242508; PubMed=12354221;
RA Glaser P., Rusniok C., Buchrieser C., Chevalier F., Prangeul L.,
RA Meadek T., Zouine M., Couve E., Lalloui L., Poyart C., Trieu-Cuot P.,
RA Kunst F.;
RT "Genome sequence of Streptococcus agalactiae, a pathogen causing
RT invasive neonatal disease.";
RL Mol. Microbiol. 45:1499-1513(2002).
RN [2]
RP SEQUENCE FROM N.A.
RX STRAIN=2603 V/R / Serotype V;
```

```
RX MEDLINE=22222988; PubMed=12200547;
RA Tettelin H., Masiugnani V., Cieslewicz M.J., Eisen J.A., Peterson S.,
RA Wessels M.R., Paulsen I.T., Nelson K.E., Margarit I., Read T.D.,
RA Madoff L.C., Wolf A.M., Beanan M.J., Brinkac L.M., Daugherty S.C.,
RA DeBoy R.T., Durkin A.S., Kolonay J.F., Madupu R., Lewis M.R.,
RA Radune D., Fedorova N.B., Scanlan D., Khouri H., Mulligan S.,
RA Carty H.A., Cline R.T., Van Aken S.E., Gill J., Scarselli M., Mora M.,
RA Iacobini E.T., Brettoni C., Galli G., Mariani M., Vegni F., Malone D.,
RA Rinaldo D., Rappuoli R., Telford J.L., Kasper D.L., Grandi G.,
RA Fraser C.M.;
RT "Complete genome sequence and comparative genomic analysis of an
RT emerging human pathogen, serotype V Streptococcus agalactiae.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:12391-12396(2002).
CC -1- SIMILARITY: BELONGS TO THE S16P FAMILY OF RIBOSOMAL PROTEINS.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announcement/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; AL766850; CAD47087.1; -
DR EMBL; AE014254; AAN00229.1; -
DR Sagalists; gbs1428; -
DR TIGR; SAG1358; -
DR HAMAP; MF 00385; -; 1.
DR InterPro; IPR000307; Ribosomal_S16.
DR Pfam; PF00886; Ribosomal_S16; 1.
DR ProDom; PD003791; Ribosomal_S16; 1.
DR TIGRFAMs; TIGR00002; S16; 1.
DR PROSITE; PS00732; RIBOSOMAL_S16; FALSE_NEG.
KW Ribosomal protein; Complete proteome.
SQ SEQUENCE 90 AA; 10282 MW; CDCA763D78007CBE CRC64;

Query Match 22.7%; Score 34.5; DB 1; Length 90;
Best Local Similarity 35.0%; Pred. No. 3.1e+02;
Matches 7; Conservative 8; Mismatches 4; Indels 1; Gaps 1;

QY 5 NPLVLMANNIIS-KEKKEIKWI 23
Dd 41 NPLVAENQVTKKEERVELEWL 60

RESULT 24
SECE_BACLI STANDARD; PRT; 59 AA.
AC P38381;
DT 01-OCT-1994 (Rel. 30, Created)
DT 01-OCT-1994 (Rel. 30, Last sequence update)
DT 01-OCT-1994 (Rel. 30, Last annotation update)
DE Preprotein translocase secE subunit.
GN SECE.
OS Bacillus licheniformis.
OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.
OX NCBI_TaxID=1402;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=88139159; PubMed=3277943;
RA Dubnau E., Weir J., Nair G., Carter L. III, Moran C.P. Jr., Smith I.;
RT "Bacillus sporulation gene spoOH codes for sigma 30 (sigma H).";
RL J. Bacteriol. 170:1054-1062(1988).
CC -1- FUNCTION: ESSENTIAL FOR PROTEIN EXPORT.
CC -1- SUBCELLULAR LOCATION: Tail-anchored membrane protein (Potential).
CC -1- SIMILARITY: BELONGS TO THE SECE/SEC61-gamma family.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announcement/
```



```
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; M29694; -; NOT ANNOTATED_CDS.
DR InterPro; IPR001901; SecE.
DR InterPro; IPR005807; SecE_bac.
DR Pfam; PF00584; SecE; 1.
DR TIGRFAMs; TIGR00964; 3a0501806; 1.
DR PROSITE; PS01067; SECE SEC61G; 1.
KW Protein transport; Translocation; Transmembrane.
FT TRANSMEM 30 50 POTENTIAL.
SQ SEQUENCE 59 AA; 6775 MW; BD40479D9FA5837B CRC64;

Query Match 22.4%; Score 34; DB 1; Length 59;
Best Local Similarity 21.1%; Pred. No. 2.4e+02;
Matches 4; Conservative 7; Mismatches 8; Indels 0; Gaps 0;

QY 4 LNVLMAMNIIISKEKEIKW 22
Db 1 MGIKFLKNVGKEMKVTW 19

RESULT 25
Y32 MARPO
ID_Y32 MARPO STANDARD; PRT; 62 AA.
AC P38474; 1994 (Rel. 30, Created)
DT 01-OCT-1994 (Rel. 30, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Hypothetical 7.2 kDa protein in ATPA-COX1 intergenic region (ORF 62).
GN YMF32.
OS Marchantia polymorpha (Liverwort).
OC Eukaryota; Viridiplantae; Streptophyta; Marchantiophyta;
OC Marchantiopsida; Marchantiales; Marchantiaceae;
OC Marchantiaceae; Marchantia.
OX NCBI_TaxID=3197;
[1]
RN J. Mol. Biol. 223:1-7 (1992).
RP SEQUENCE FROM N.A. PubMed=1731062;
RX MEDLINE=92114051; Ohta K., Nakamura Y., Takemura M., Nozato N.,
RA Akashi K., Kanegae T., Ogura Y., Kohchi T., Ohshima K.;
RT "Gene organization deduced from the complete sequence of liverwort
RT Marchantia polymorpha mitochondrial DNA. A primitive form of plant
RT mitochondrial genome."
RL J. Mol. Biol. 223:1-7 (1992).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; M68929; AAC09448.1; -.
DR PIR; S25999; S25999.
KW Mitochondrion; Hypothetical protein.
SQ SEQUENCE 62 AA; 7231 MW; D98D55A9EF768E64 CRC64;

Query Match 22.4%; Score 34; DB 1; Length 62;
Best Local Similarity 28.6%; Pred. No. 2.5e+02;
Matches 6; Conservative 5; Mismatches 10; Indels 0; Gaps 0;

QY 4 LNVLMAMNIIISKEKEIKW 24
Db 8 LNTFRQLTLPKNGNTFVG 28

RESULT 26
CSPA_RICCN
ID_CSPA_RICCN STANDARD; PRT; 70 AA.
AC Q92GV1;
DT 28-FEB-2003 (Rel. 41, Created)

or send an email to license@isb-sib.ch).
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Cold shock-like protein csppA.
GN CSPPA OR RCL021.
OS Rickettsia conorii.
OC Bacteria; Proteobacteria; Alphaproteobacteria; Rickettsiales;
OC Rickettsiaceae; Rickettsiae; Rickettsia.
OX NCBI_TaxID=781;
[1]
RN SEQUENCE FROM N.A.
RC STRAIN=Malish 7;
RX MEDLINE=21442074; PubMed=11557893;
RA Ogata H., Audic S., Renesto-Audifern P., Fournier P.-E., Barbe V.,
RA Samson D., Roux V., Cossart P., Weissbach J., Claverie J.-M.,
RA Raoult D.;
RT "Mechanisms of evolution in Rickettsia conorii and R. prowazekii.";
Science 293:2093-2098 (2001).
RL -!- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
CC -!- SIMILARITY: BELONGS TO THE COLD-SHOCK DOMAIN (CSD) FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; AE008653; AL03559.1; -.
DR PIR; E97827; E97827.
DR InterPro; IPR002059; Cold_shock.
DR Pfam; PF00313; CSD; 1.
DR PRINTS; PR00050; COLDSHOCK.
DR ProDom; PD000621; Cold shock; 1.
DR SMART; SM00357; CSP; 1.
DR PROSITE; PS00352; COLD_SHOCK; 1.
KW Transcription regulation; DNA-binding; Activator; Complete proteome.
FT DOMAIN 7 67 CSD.
SQ SEQUENCE 70 AA; 7771 MW; 822E9714229FE03A CRC64;

Query Match 22.4%; Score 34; DB 1; Length 70;
Best Local Similarity 38.1%; Pred. No. 2.8e+02;
Matches 8; Conservative 2; Mismatches 7; Indels 4; Gaps 1;

QY 8 MAMNIIISKEKEIKWIGLPTN 28
Db 1 MATNIVGK----VKWYNSTKN 17

RESULT 27
CSPPA_RICPR
ID_CSPA_RICPR STANDARD; PRT; 70 AA.
AC Q92CP9;
DT 30-MAY-2000 (Rel. 39, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Cold shock-like protein csppA.
GN CSPPA OR RP670.
OS Rickettsia prowazekii.
OC Bacteria; Proteobacteria; Alphaproteobacteria; Rickettsiales;
OC Rickettsiaceae; Rickettsiae; Rickettsia.
OX NCBI_TaxID=782;
[1]
RN SEQUENCE FROM N.A.
RC STRAIN=Madrid E;
RX MEDLINE=95039499; PubMed=9823893;
RA Andersson S.G.E., Zomorodipour A., Andersson J.O.,
RA Sichertitz-Ponten T., Alsmark U.C.M., Podowski R.M., Naeslund A.K.,
RA Eriksson A.-S., Winkler H.H., Kurland C.G.;
RT "The genome sequence of Rickettsia prowazekii and the origin of
RT mitochondria."
RL Nature 396:133-140 (1998).
RL -!- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
CC -!- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
CC -----
```



CC -1- SIMILARITY: BELONGS TO THE COLD-SHOCK DOMAIN (CSD) FAMILY.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (see <http://www.isb-sib.ch/announcement/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----

DR EMBL; AJ235272; CAA15108.1; --  
DR PIR; B71673; B71673.  
DR HSP; P41016; IC90.  
DR InterPro; IPR002059; Cold\_shock.  
DR Pfam; PF00313; CSD; 1.  
DR PRINTS; PD00050; COLDSHOCK.  
DR ProDom; PD000621; Cold\_shock; 1.  
DR SMART; SM00357; CSP; 1.  
DR PROSITE; PS00352; COLD\_SHOCK; 1.  
KW Transcription regulation; DNA-binding; Activator; Complete proteome.  
FT DOMAIN 7 67 CSD.  
SQ SEQUENCE 70 AA; 7785 MW; F49FBF2FB344903E CRC64;

Query Match 22.4%; Score 34; DB 1; Length 70;  
Best Local Similarity 38.1%; Pred. No. 2.8e+02;  
Matches 8; Conservative 2; Mismatches 7; Indels 4; Gaps 1;

QY 8 MAMNIISKEKEIKWGLPTN 28  
DB 1 MATNIVGK----VKWYNSTKN 17

RESULT 28

RL15\_BACLI STANDARD; PRT; 72 AA.  
AC P35136;  
DT 01-FEB-1994 (Rel. 28, Created)  
DT 01-FEB-1994 (Rel. 28, Last sequence update)  
DT 01-FEB-1994 (Rel. 28, Last annotation update)  
DE 50S ribosomal protein L15 (Fragment).  
GN RPL0.  
OS Bacillus licheniformis.  
OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.  
OX NCBI\_TaxID=1402;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=93062802; PubMed=1435726;  
RA Tschander S., Driessen A.J.M., Freud R.;  
RT "Cloning and molecular characterization of the secY genes from  
RT Bacillus licheniformis and Staphylococcus carnosus: comparative  
RT analysis of nine members of the SecY family.";  
RL Mol. Gen. Genet. 235:147-152(1992).  
CC -1- FUNCTION: THIS PROTEIN BINDS DIRECTLY TO 23S RIBOSOMAL RNA.  
CC -1- SIMILARITY: BELONGS TO THE L15 FAMILY OF RIBOSOMAL PROTEINS.  
CC -----

CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (see <http://www.isb-sib.ch/announcement/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----

DR EMBL; X70087; -- NOT\_ANNOTATED\_CDS.  
DR PIR; S34404; S34404.  
DR InterPro; IPR001196; Ribosomal\_L15.  
DR Pfam; PF00256; L15; 1.  
DR Pfam; PF01305; Ribosomal\_L15; 1.  
DR PROSITE; PS00475; RIBOSOMAL\_L15; 1.  
KW Ribosomal protein; rRNA-binding.  
FT NON\_TER 1  
SQ SEQUENCE 72 AA; 7441 MW; 6DC3BA681492E6C9 CRC64;

Query Match 22.4%; Score 34; DB 1; Length 72;  
Best Local Similarity 36.8%; Pred. No. 2.9e+02;  
Matches 7; Conservative 5; Mismatches 7; Indels 0; Gaps 0;

QY 6 VLMAMNIISKEKEIKWIG 24  
DB 21 LLELTGVISKRGVKILG 39

RESULT 29

LEU2\_HUMAN STANDARD; PRT; 84 AA.  
ID LEU2\_HUMAN  
AC O43262;  
DT 16-OCT-2001 (Rel. 40, Created)  
DT 16-OCT-2001 (Rel. 40, Last sequence update)  
DT 28-FEB-2003 (Rel. 41, Last annotation update)  
DE Leukemia associated protein 2 (Deleted in lymphocytic leukemia 2).  
GN DLEU2 OR LEU2.  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
OC Mammalia; Euthera; Primates; Catarrhini; Homnidae; Homo.  
OX NCBI\_TaxID=9606;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=98055620; PubMed=9395242;  
RA Liu Y., Corcoran M., Rasool O., Ivanova G., Ibbotson R., Grander D.,  
RA Iyengar A., Baranova A., Kashuba V., Merup M., Wu X., Gardiner A.,  
RA Mullenbach R., Poltarau A., Hultstrom A.L., Juliusson G., Chapman R.,  
RA Tiller M., Cotter F., Gahrton G., Yankovsky N., Zabarovsky E.,  
RA Einhorn S., Oscier D.;  
RT "Cloning of two candidate tumor suppressor genes within a 10 kb region  
RT on chromosome 13q14, frequently deleted in chronic lymphocytic  
RT leukemia.";  
RL Oncogene 15:2463-2473(1997).  
CC -1- FUNCTION: MAY ACT AS A TUMOR SUPPRESSOR.  
CC -----

CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (see <http://www.isb-sib.ch/announcement/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----

DR EMBL; Y15228; CAA75516.1; --  
DR Genew; HGNC:13748; DLEU2.  
DR MIM; 605766; --  
DR GO; GO:0008181; F:tumor suppressor; TAS.  
KW Anti-oncogene.  
SQ SEQUENCE 84 AA; 10196 MW; D1EE021072A0E03B CRC64;

Query Match 22.4%; Score 34; DB 1; Length 84;  
Best Local Similarity 46.7%; Pred. No. 3.4e+02;  
Matches 7; Conservative 2; Mismatches 6; Indels 0; Gaps 0;

QY 4 LNVLMAMNIISKEKK 18  
DB 68 LKCYFSLNTIKKK 82

RESULT 30

FCEG\_RAT STANDARD; PRT; 86 AA.  
ID FCEG\_RAT  
AC P20411;  
DT 01-FEB-1991 (Rel. 17, Created)  
DT 01-FEB-1991 (Rel. 17, Last sequence update)  
DT 28-FEB-2003 (Rel. 41, Last annotation update)  
DE High affinity immunoglobulin epsilon receptor gamma-subunit precursor  
DE (FCER1) (Ige Fc receptor gamma-subunit) (fc-epsilon RI-gamma).  
GN FCER1G OR FCE1G.  
OS Rattus norvegicus (Rat).  
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
OC Mammalia; Euthera; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.

```

Query Match      22.4%; Score 34; DB 1; Length 95;
Best Local Similarity 36.8%; Pred. NO. 3.9e+02;
Matches 7; Conservative 6; Mismatches 6; Indels 0; Gaps 0;

QY      5 NVLMAMNIISKKEIKKIWI 23
       |||::|||::|||::|||
Db       37 NLLKIQSIKKKKLTKFI 55

RESULT 33
Y117 NPVAC
ID _Y117 NPVAC STANDARD; PRT; 95 AA.
```

```

AC P41670;
DT 01-NOV-1995 (Rel. 32, Created)
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DE 01-NOV-1997 (Rel. 35, Last annotation update)
DE Hypothetical 11.0 kDa protein in HB65-PK2 intergenic region.
OS Autographa californica nuclear polyhedrosis virus (AcMNPV).
OC Viruses; dsDNA viruses, no RNA stage; Baculoviridae;
OC Nucleopolyhedrovirus.
OX NCBI_TaxID=46015;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C6;
RX MEDLINE=94303173; PubMed=8030224;
RA Ayres M.D., Howard S.C., Kuzio J., Lopez-Ferber M., Possee R.D.;
RT "The complete DNA sequence of Autographa californica nuclear
RT polyhedrosis virus."
RL Virology 202:586-605(1994).
CC -1- SIMILARITY: TO CORRESPONDING ORF IN OPMNPV.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; L22858; AAA66747.1; -.
DR EMBL; F72864; F72864.
KW Hypothetical protein.
SQ SEQUENCE 95 AA; 10992 MW; 68897B745ECB51C6 CRC64;

Query Match 22.4%; Score 34; DB 1; Length 95;
Best Local Similarity 36.8%; Pred. No. 3.9e+02;
Matches 7; Conservative 3; Mismatches 3; Indels 2; Gaps 1;

QY 5 NVLMAMNIIISKEKKEIKWI 23
Db 6 NVLLVPNAL--KKRDVKYI 22

RESULT 34
SPAC_BPT4 STANDARD; PRT; 97 AA.
AC P39230;
DT 01-FEB-1995 (Rel. 31, Created)
DT 01-FEB-1995 (Rel. 31, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Spackie protein precursor.
GN SP OR 61.3.
OS Bacteriophage T4.
OC Viruses; dsDNA viruses, no RNA stage; Caudovirales; Myoviridae;
OC T4-like viruses
OC NCBI_TaxID=10665;
OX [1]
RN SEQUENCE FROM N.A.
RX MEDLINE=93188183; PubMed=8383243;
RA Selick H.E., Stormo G.D., Dyson R.L., Alberts B.M.;
RT "Analysis of five presumptive protein-coding sequences clustered
RT between the primosome genes, 41 and 61, of bacteriophages T4, T2, and
RT T6."
RL J. Virol. 67:2305-2316(1993).
RN [2]
RP SEQUENCE FROM N.A.
RA Kutter E., Arisaka F., Kunisawa T., Tsugita A., Mosig G.,
RA Mesyanzhinov V., Ruger W., Stidham T., Thomas E.;
RT "Bacteriophage T4 genome analysis."
RL Submitted (JUL-2000) to the EMBL/GenBank/DBJ databases.
RN [3]
RP CHARACTERIZATION
RX MEDLINE=99348509; PubMed=10417260;
RA Kai T., Ueno H., Otsuka Y., Morimoto W., Yonesaki T.;
RT "Gene 61.3 of bacteriophage T4 is the spackie gene."

```

```

RL Virology 260:254-259(1999).
CC -1- FUNCTION: NOT KNOWN. MUTANTS ALLOW SURVIVAL OF LYSOZYME MUTANTS.
CC -1- SUBCELLULAR LOCATION: Periplasmic (Potential).
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; S57514; AAB25710.1; -.
DR EMBL; AF158101; AAD42510.1; -.
DR PIR; C45681; C45681.
KW Signal; Periplasmic.
FT SIGNAL 1 22 POTENTIAL.
FT CHAIN 23 97 SPACKLE PROTEIN.
SQ SEQUENCE 97 AA; 10994 MW; E0A5E5E076C97965 CRC64;

Query Match 22.4%; Score 34; DB 1; Length 97;
Best Local Similarity 41.7%; Pred. No. 3.9e+02;
Matches 5; Conservative 5; Mismatches 2; Indels 0; Gaps 0;

QY 10 MNIIISKEKKEIK 21
Db 49 MNIVKRDPRPMK 60

RESULT 35
RET5_BOVIN STANDARD; PRT; 42 AA.
AC P82708;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Retinol-binding protein III, cellular (CRBP-III) (Fragment).
GN RBP5.
OS Bos taurus (Bovine).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
OC Bovidae; Bovinae; Bos.
OX NCBI_TaxID=9913;
RN [1]
RP SEQUENCE, AND FUNCTION.
RC TISSUE=Kidney;
RX MEDLINE=21173623; PubMed=11274389;
RA Follis C., Calderone V., Ottonello S., Bolchi A., Zanotti G.,
RA Stoppani M., Berni R.;
RT "Identification, retinoid binding and X-ray analysis of a human
RT retinol-binding protein."
RL Proc. Natl. Acad. Sci. U.S.A. 98:3710-3715(2001).
CC -1- FUNCTION: Intracellular transport of retinol.
CC -1- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
CC -1- TISSUE SPECIFICITY: Kidney.
CC -1- SIMILARITY: BELONGS TO THE FABP/P2/CRBP/CRABP FAMILY OF
CC TRANSPORTERS.
DR HSP; P82980; IGLL.
DR InterPro; IPR000463; Fatty acid BP.
DR InterPro; IPR000566; Lipocalin_cyFABP.
DR Pfam; PF00061; Lipocalin; 1.
DR PROSITE; PS00214; FABP; FALSE NEG.
KW Vitamin A; Retinol-binding; Transport.
FT NON TER 42
SQ SEQUENCE 42 AA; 4892 MW; ACH4F1399FDD7F09 CRC64;

Query Match 22.0%; Score 33.5; DB 1; Length 42;
Best Local Similarity 47.6%; Pred. No. 2e+02;
Matches 10; Conservative 3; Mismatches 5; Indels 3; Gaps 1;

QY 3 ALNVLAMN---IISKEKKEI 20
Db 22 ALNVLAMNRIALLKPKDEI 42

```

5 NVLMAMNIIIS-KEKKEIKWI 23

Db 41 NPLVAENQITIKEDRVLEWL 60

RESULT 38  
RS16 STRMU STANDARD; PRT; 91 AA.

AC QSDUN9;  
DT 15-SEP-2003 (Rel. 42, Created)  
DT 15-SEP-2003 (Rel. 42, Last sequence update)  
DT 15-SEP-2003 (Rel. 42, Last annotation update)  
DE 308 ribosomal protein S16.  
GN RPSL OR SMU.865.  
OS Streptococcus mutans.  
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;  
OC Streptococcus.  
OX NCBI\_TaxID=1309;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=UAI59 / ATCC 700610 / Serotype C;  
RX MEDLINE=22295063; PubMed=12397186;  
RA Ajdic D., McShan W.M., McLaughlin R.E., Savic G., Chang J.,  
RA Carson M.B., Primeaux C., Tian R., Kenton S., Jia H., Lin S., Qian Y.,  
RA Li S., Zhu H., Najjar F., Lai H., White J., Roe B.A., Ferretti J.J.;  
RT "Genome sequence of Streptococcus mutans UAI59, a cariogenic dental  
RT pathogen.";  
RL Proc. Natl. Acad. Sci. U.S.A. 99:14434-14439(2002).  
CC -1- SIMILARITY: BELONGS TO THE S16P FAMILY OF RIBOSOMAL PROTEINS.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
CC DR EMBL; AE014528; AAN58580.1; -  
DR HAMAP; MF 00385; -; 1.  
DR InterPro; IPR000307; Ribosomal\_S16.  
DR Pfam; PF00886; Ribosomal\_S16; 1.  
DR ProDom; PD003791; Ribosomal\_S16; 1.  
DR TIGRFAMs; TIGR00002; S16; 1.  
DR PROSITE; PS00732; RIBOSOMAL\_S16; FALSE\_NEG.  
KW Ribosomal protein; Complete proteome.  
SQ SEQUENCE 91 AA; 10410 MW; 1E7329D1BDFF6EC2 CRC64;  
Query Match 22.08; Score 33.5; DB 1; Length 91;  
Best Local Similarity 33.34; Pred. No. 4.4e+02;  
Matches 7; Conservative 7; Mismatches 6; Indels 1; Gaps 1;  
Qy 5 NVLMAMNIIIS-KKKKEIKWIG 24  
Db 41 NPLVTENQVTKEDRVLEWL 61

RESULT 39  
RUXE YEAST STANDARD; PRT; 94 AA.

AC Q12330;  
DT 01-NOV-1997 (Rel. 35, Created)  
DT 01-NOV-1997 (Rel. 35, Last sequence update)  
DT 15-SEP-2003 (Rel. 42, Last annotation update)  
DE Small nuclear ribonucleoprotein E (snRNP-E) (Sm protein E) (Sm-E)  
DE (SME).  
GN SME1 OR YOR159C.  
OS Saccharomyces cerevisiae (Baker's yeast).  
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;  
OC Saccharomycetales; Saccharomycetaceae; Saccharomycetes.  
OX NCBI\_TaxID=4932;  
RN [1]  
RP SEQUENCE FROM N.A., AND FUNCTION.  
RC STRAIN=S288C;  
RX MEDLINE=97075918; PubMed=8918241;

RA Bordonne R., Tarasov I.A.;  
RT "The yeast SME1 gene encodes the homologue of the human E core  
RL protein.";  
RL Gene 176:111-117(1996).  
RN [2]  
RP SEQUENCE FROM N.A.  
RC STRAIN=S288C / FY1678;  
RA Ayadi A., Bordonne R., Camasses A., Madania A., Poch O.,  
RA Tarasov I.A., Winsor B., Martin R.P.;  
RL Submitted (MAY-1996) to the EMBL/GenBank/DBJ databases.  
RN [3]  
RP SEQUENCE FROM N.A., AND INTERACTION WITH SMX2 AND SMX3.  
RX MEDLINE=98187616; PubMed=9528767;  
RA Camasses A., Bragado-Nilsson E., Martin R., Seraphin B., Bordonne R.;  
RT "Interactions within the yeast Sm core complex: from proteins to amino  
RL acids.";  
RL Mol. Cell. Biol. 18:1956-1966(1998).  
RN [4]  
RP SEQUENCE FROM N.A., AND RNA-BINDING.  
RX MEDLINE=99298196; PubMed=10369684;  
RA Salgado-Garrido J., Bragado-Nilsson E., Kandel-Lewis S., Seraphin B.;  
RT "Sm and Sm-like proteins assemble in two related complexes of deep  
RL evolutionary origin.";  
RL EMBO J. 18:3451-3462(1999).  
RN [5]  
RP PARTIAL SEQUENCE.  
RX MEDLINE=97165029; PubMed=9012791;  
RA Neubauer G., Gottschalk A., Fabrizio P., Seraphin B., Luehrmann R.,  
RA Mann M.;  
RT "Identification of the proteins of the yeast U1 small nuclear  
RT ribonucleoprotein complex by mass spectrometry.";  
RL Proc. Natl. Acad. Sci. U.S.A. 94:385-390(1997).  
RN [6]  
RP CHARACTERIZATION OF THE SPLICEOSOME.  
RX MEDLINE=21664121; PubMed=11804584;  
RA Stevens S.W., Ryan D.E., Ge H.Y., Moore R.E., Young M.K., Lee T.D.,  
RA Abelson J.;  
RT "Composition and functional characterization of the yeast spliceosomal  
RL penta-snRNP.";  
RL Mol. Cell 9:31-44(2002).  
RN [7]  
RP SUBUNITS.  
RX MEDLINE=21199560; PubMed=11302706;  
RA Walke S., Bragado-Nilsson E., Seraphin B., Nagai K.;  
RT "Stoichiometry of the Sm proteins in yeast spliceosomal snRNPs  
RT supports the heptamer ring model of the core domain.";  
RL J. Mol. Biol. 308:49-58(2001).  
RN [8]  
RP SUBCELLULAR LOCATION.  
RX MEDLINE=20481754; PubMed=11027265;  
RA Bordonne R.;  
RT "Functional characterization of nuclear localization signals in yeast  
RL Sm proteins.";  
RL Mol. Cell. Biol. 20:7943-7954(2000).  
CC -1- FUNCTION: Involved in pre-mRNA splicing. Binds and is required for  
CC the stability of sn-RNA U1, U2, U4 and U5 which contain a highly  
CC conserved structural motif called the Sm binding site. Involved in  
CC cap modification.  
CC -1- SUBUNIT: Component of the Sm core complex, present in spliceosomal  
CC snRNP U1, U2, U4/U6 and U5. The core complex contains SME1, SME1,  
CC SME2, SME3, SME1, SME3 and SMX2 (Sm proteins B, D1, D2, D3, E, F  
CC and G, respectively), and is probably a heptameric ring structure.  
CC SME1 specifically interacts with SMX2 and SMX3.  
CC -1- SUBCELLULAR LOCATION: Nuclear and cytoplasmic.  
CC -1- SIMILARITY: Belongs to the snRNP Sm proteins family.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----

```

CC EMBL; X32449; CAA63198.1; -
DR EMBL; U55020; AAC49645.1; -
DR EMBL; Z75067; CAA99365.1; -
DR PIR; S67047; S67047.
DR SGD; S0005685; SMEI.
DR GO; GO:0030532; C-small nuclear ribonucleoprotein complex; IDA.
DR GO; GO:0008248; P-pre-mRNA splicing factor activity; IPI.
DR GO; GO:0006371; P-mRNA splicing; IPI.
DR InterPro; IPR001163; snRNP_Sm.
DR Pfam; PF01423; Sm; 1.
DR SMART; SM00651; Sm; 1.
KW Nuclear protein; Ribonucleoprotein; mRNA splicing; mRNA processing;
RNA-binding.
SQ SEQUENCE 94 AA; 10375 MW; B76D72DA8456F3B5 CRC64;

Query Match 22.0%; Score 33.5; DB 1; Length 94;
Best Local Similarity 42.9%; Pred. No. 4.5e+02;
Matches 9; Conservative 2; Mismatches 5; Indels 5; Gaps 1;

Qy 10 MNIIISKEKEIKWGLPTNSA 30
Db 50 MNVVIDEAVEI-----PVNSA 65

RESULT 40
RL32 MAIZE STANDARD; PRT; 42 AA.
AC P51421.
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 01-OCT-1996 (Rel. 34, Last annotation update)
DE 60S ribosomal protein L32 (Fragment).
GN RPL32.
OS Zea mays (Maize).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC PACCAD clade; Panicoideae; Andropogoneae; Zea.
OX NCBI_TaxID=4577;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CV. HD5 X HD7;
RA Bates E.E.M., Vergne P., Dumas C.;
RA Submitted (NOV-1993) to the EMBL/GenBank/DBJ databases.
RL -!- SIMILARITY: BELONGS TO THE L32E FAMILY OF RIBOSOMAL PROTEINS.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; X75646; CAA53301.1; -
DR PIR; S38633; S38633.
DR MaizeDB; 61651.
DR InterPro; IPR001515; Ribosomal L32E.
DR Pfam; PF01655; Ribosomal L32e; 1.
DR ProDom; PD003823; Ribosomal L32E; 1.
DR PROSITE; PS00580; RIBOSOMAL_L32E; PARTIAL.
KW Ribosomal protein.
FT NON TER 1
FT SEQUENCE 42 AA; 4709 MW; FA9E004E030D1BF4 CRC64;

Query Match 21.7%; Score 33; DB 1; Length 42;
Best Local Similarity 46.2%; Pred. No. 2.4e+02;
Matches 6; Conservative 5; Mismatches 2; Indels 0; Gaps 0;

Qy 8 MANNIIISKEKEI 20
Db 6 IAHNVSTKKKEI 18

```

```

RESULT 41
SECE_BACSU STANDARD; PRT; 59 AA.
ID SECE_BACSU
AC Q06799; P36689;
DT 01-JUN-1994 (Rel. 29, Created)
DT 01-JUN-1994 (Rel. 29, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Preprotein translocase secE subunit.
GN SECE.
OS Bacillus subtilis.
OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.
OX NCBI_TaxID=1423;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=95058172; PubMed=7968510;
RA Jeong S., Yoshikawa H., Takahashi H.;
RT "Isolation and characterization of the secE homologue gene of
RT Bacillus subtilis.";
RL Mol. Microbiol. 10:133-142 (1993).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=168;
RX MEDLINE=98044033; PubMed=9384377;
RA Kunat F., Ogasawara N., Moszer I., Albertini A.M., Alloni G.,
RA Azevedo V., Bertero M.G., Bessieres P., Bolotin A., Borchert S.,
RA Boriss R., Boursier L., Brans A., Braun M., Brignell S.C., Bron S.,
RA Brouillet S., Bruschi C.V., Caldwell B., Capuano V., Carter N.M.,
RA Choi S.K., Codani J.J., Connerton I.F., Cummings N.J., Daniel R.A.,
RA Denizot F., Devine K.M., Dusterhoft A., Ehrlich S.D., Emmerson P.T.,
RA Entian K.D., Errington J., Fabret C., Ferrari E., Foulger D.,
RA Fritz C., Fujita M., Fujita Y., Fuma S., Galizzi A., Galleron N.,
RA Ghim S.Y., Glaser P., Goffeau A., Gollightly E.J., Grandi G., Henaut A.,
RA Guisepi G., Guy B.J., Haga K., Halech J., Harwood C.R., Jones L.,
RA Hilbert H., Holsappel S., Hosono S., Hullo M.F., Itaya M., Klein C.,
RA Joris B., Karamata D., Kasahara Y., Kjaer-Blanchard M., Kumano M.,
RA Kobayashi Y., Koetter P., Koningsstein G., Krogh S., Lazarevic V.,
RA Kurita K., Lapidus A., Lardinois S., Lauber J., Lazarevic V.,
RA Lee S.M., Levine A., Liu H., Masuda S., Maul C., Medigue C.,
RA Medina N., Mellado R.P., Mizuno M., Moesti D., Nakai S., Noback M.,
RA Noone D., O'Reilly M., Ogawa K., Ogiwara A., Oudega B., Park S.H.,
RA Parro V., Pohl T.M., Portetelle D., Porwollik S., Prescott A.M.,
RA Plescan E., Pujic P., Purnelle B., Rapoport G., Rey M., Reynolds S.,
RA Rieger M., Rivolta C., Rocha E., Roche B., Rose M., Sadale Y.,
RA Sato T., Scanlan E., Schleich S., Schroeter R., Scoffone F.,
RA Sekiguchi J., Sekowska A., Seror S.J., Serror P., Shin B.S., Soldo B.,
RA Sorokin A., Tacconi E., Takagi T., Takahashi H., Takemaru K.,
RA Tasechi M., Tamakoshi A., Tanaka T., Terpestra P., Tognoni A.,
RA Tesato V., Uchiyama S., Vandenbol M., Vannier F., Vassarotti A.,
RA Viari A., Wambutt R., Wedler E., Wedler H., Weitzenegger T.,
RA Winters P., Wipat A., Yamamoto H., Yamane K., Yasumoto K., Yata K.,
RA Yoshida K., Yoshikawa H.F., Zumstein E., Yoshikawa H., Danchin A.;
RT "The complete genome sequence of the Gram-positive bacterium Bacillus
RT subtilis.";
RL Nature 390:249-256 (1997).
RL Nature 390:249-256 (1997).
CC -!- FUNCTION: ESSENTIAL FOR PROTEIN EXPORT.
CC -!- SUBCELLULAR LOCATION: Tail-anchored membrane protein (Potential).
CC -!- SIMILARITY: Belongs to the secE/SEC61-gamma family.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; D13303; BAA02559.1; -
DR EMBL; Z99104; CAB11876.1; -
DR PIR; S39858; S39858.
DR Subtilist; BG10161; secE.
DR InterPro; IPR001901; SecE.
DR InterPro; IPR005807; SecE_bac.

```





```
Db 1 MNLNVNRELKVGQIKNGVSSA 24
RESULT 44
Y12J_BPT4 STANDARD; PRT; 65 AA.
ID Y12J_BPT4
AC Q02406;
DT 01-OCT-1993 (Rel. 27, Created)
DT 01-OCT-1993 (Rel. 27, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Hypothetical 7.3 kDa protein in Gp30-rIII intergenic region.
GN Y12J OR 30.5.
OS Bacteriophage T4.
OC Viruses; dsDNA viruses, no RNA stage; Caudovirales; Myoviridae;
OC T4-like viruses.
OX NCBI_TaxID=10665;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=93076002; PubMed=1446076;
RA Nivinskas R., Zajackauskaite A., Raudonikienė A., Vitenienė I.;
RT "The nucleotide sequence between genes 31 and 30 of bacteriophage
RT T4."
RL DNA Seq. 2:405-409(1992).
RN [2]
RP SEQUENCE FROM N.A.
RA Kutter E., Arisaka F., Kunisawa T., Tsugita A., Mosig G.,
RA Mesyanzhinov V., Ruger W., Stidham T., Thomas E.;
RT "Bacteriophage T4 genome analysis."
RL Submitted (JUL-2000) to the EMBL/GenBank/DBJ databases.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.ebi.ac.uk/announcements/
CC or send an email to license@ebi.ac.uk).
CC -----
CC EMBL; X60109; CAA42705.1; -.
CC DR EMBL; AF158101; AAD42446.1; -.
CC DR PIR; S27147; S27147.
CC KW Hypothetical protein.
SQ SEQUENCE 65 AA; 7252 MW; F6F53A9D21DFB4A3 CRC64;
Query Match 21.7%; Score 33; DB 1; Length 65;
Best Local Similarity 36.8%; Pred. No. 3.7e+02;
Matches 7; Conservative 2; Mismatches 10; Indels 0; Gaps 0;
QY 12 IISKKKKEIKWIGLPTNSA 30
Db 19 LISKVLPPFKWSGIQIKEA 37
RESULT 45
VG5_SPV4 STANDARD; PRT; 83 AA.
ID VG5_SPV4
AC P11337;
DT 01-JUL-1989 (Rel. 11, Created)
DT 01-JUL-1989 (Rel. 11, Last sequence update)
DT 01-JUL-1989 (Rel. 11, Last annotation update)
DE Gene 5 protein.
GN 5.
OS Spiroplasma virus 4 (Spv4).
OC Viruses; ssDNA viruses; Microviridae; Spiromicrovirus.
OX NCBI_TaxID=10855;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=88032809; PubMed=2822658;
RA Renaudin J., Pascarel M.-C., Bove J.-M.;
RT "Spiroplasma virus 4: nucleotide sequence of the viral DNA,
RT regulatory signals, and proposed genome organization."
RL J. Bacteriol. 169:4950-4961(1987).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.ebi.ac.uk/announcements/
CC or send an email to license@ebi.ac.uk).
CC -----
CC EMBL; M17988; -; NOT ANNOTATED_CDS.
CC DR EMBL; D29825; G5BPSV_
SQ SEQUENCE 83 AA; 9494 MW; F6E678B3CCF75A25 CRC64;
Query Match 21.7%; Score 33; DB 1; Length 83;
Best Local Similarity 66.7%; Pred. No. 4.8e+02;
Matches 4; Conservative 2; Mismatches 0; Indels 0; Gaps 0;
QY 21 KWIGLP 26
Db 35 KWLGIP 40
Search completed: February 17, 2004, 10:57:04
Job time : 9.50495 secs
```



GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: February 17, 2004, 10:50:13 ; Search time 18.1188 Seconds  
(without alignments)  
159.230 Million cell updates/sec

Title: US-09-900-147-6  
Perfect score: 152  
Sequence: 1 YDALNVLMAMNIISKEKEIKWIGLPTNSA 30  
Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 283308 seqs, 96168682 residues  
Total number of hits satisfying chosen parameters: 37673

Minimum DB seq length: 0  
Maximum DB seq length: 100

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : PIR 76.\*  
1: pir1.\*  
2: pir2.\*  
3: pir3.\*  
4: pir4.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	43	28.3	78	2	G90530
2	42	27.6	61	2	I59337
3	42	27.6	74	2	T37447
4	42	27.6	76	2	H90332
5	42	27.6	96	2	B89941
6	40	26.3	69	2	S66650
7	40	26.3	76	2	B90299
8	40	26.3	76	2	B90331
9	40	26.3	90	2	B81102
10	39	25.7	64	2	D69406
11	39	25.7	80	2	D97909
12	39	25.7	92	2	AG1511
13	38.5	25.3	45	2	D81523
14	38.5	25.3	99	1	G64330
15	38	25.0	72	2	JQ1805
16	38	25.0	88	2	A42527
17	38	25.0	91	1	C64069
18	38	25.0	92	2	AF1152
19	37.5	24.7	83	2	H40888
20	37.5	24.7	95	2	T43125
21	37	24.3	84	2	T47420
22	37	24.3	87	2	B41869
23	37	24.3	88	2	H90329
24	37	24.3	96	2	E83815
25	36.5	24.0	69	2	T45696
26	36	23.7	42	2	T07291
27	36	23.7	44	2	A89995
28	36	23.7	56	2	B64368
29	36	23.7	60	2	JC1386

30	36	23.7	67	2	AE3378	hypothetical prote
31	36	23.7	73	2	T48368	hypothetical prote
32	36	23.7	74	2	G69286	hypothetical prote
33	36	23.7	75	2	A24500	Real hypothetical
34	36	23.7	77	2	S33665	hypothetical prote
35	36	23.7	78	2	JQ1781	Salp8R 8.8K protei
36	36	23.7	79	2	AC1217	hypothetical prote
37	36	23.7	80	2	AD1134	hypothetical prote
38	36	23.7	88	2	B64888	YnaE protein - Esc
39	36	23.7	88	2	C64909	YnfK protein - Esc
40	36	23.7	96	2	E69951	dihydrodipicolinat
41	36	23.7	97	2	F65108	hypothetical 10.8
42	36	23.7	97	2	C91136	hypothetical prote
43	36	23.7	97	2	F85981	hypothetical prote
44	36	23.7	97	2	H71158	hypothetical prote
45	36	23.7	98	2	B95983	hypothetical prote

ALIGNMENTS

RESULT 1

G90530

hypothetical protein MYPV\_1510 [imported] - Mycoplasma pulmonis (strain UAB CTIP)  
C;Species: Mycoplasma pulmonis  
C;Date: 24-May-2001 #sequence\_revision 24-May-2001 #text\_change 03-Aug-2001  
C;Accession: G90530  
R;Chambaud, I.; Heilig, R.; Ferris, S.; Barbe, V.; Samson, D.; Galissou, F.; Moszer, I.;  
Nucleic Acids Res. 29, 2145-2153, 2001  
A;Title: The complete genome sequence of the murine respiratory pathogen Mycoplasma pulmonis  
A;Reference number: A99512; MUID:21267165; PMID:11353084  
A;Accession: G90530  
A;Status: preliminary  
A;Molecule type: DNA  
A;Residues: 1-78 <KUR>  
A;Cross-references: GB:AL445566; PID:G14089564; PIDN:CAC13324.1; GSPDB:GN00153  
A;Experimental source: strain UAB CTIP  
C;Genetics:  
A;Gene: MYPV\_1510  
A;Genetic code: SGC3

Query Match 28.3%; Score 43; DB 2; Length 78;  
Best Local Similarity 45.0%; Pred. No. 33;  
Matches 9; Conservative 5; Mismatches 6; Indels 0; Gaps 0;

Qy 4 LNVLMAMNIISKEKEIKWI 23  
||| : : : ||| : : ||  
Db 9 LNVKLIDHISKKNFKDEWI 28

RESULT 2

I59337

mammary transforming protein - mouse  
C;Species: Mus musculus (house mouse)  
C;Date: 02-Jul-1996 #sequence\_revision 02-Jul-1996 #text\_change 05-Nov-1999  
C;Accession: I59337  
R;Bera, T.K.; Guman, R.C.; Miyamoto, S.; Panda, D.K.; Sasaki, M.; Hanyu, K.; Enami, J.;  
Proc. Natl. Acad. Sci. U.S.A. 91, 9789-9793, 1994  
A;Title: Identification of a mammary transforming gene (MAT1) associated with mouse mammary  
A;Reference number: I59337; MUID:95024046; PMID:7937892  
A;Accession: I59337  
A;Status: preliminary; translated from GB/EMBL/DBJ  
A;Molecule type: mRNA  
A;Residues: 1-61 <RES>  
A;Cross-references: GB:U31958; NID:G473909; PIDN:AAC37665.1; PID:G473910  
C;Genetics:  
A;Gene: MAT1

Query Match 27.6%; Score 42; DB 2; Length 61;  
Best Local Similarity 33.3%; Pred. No. 35;  
Matches 7; Conservative 9; Mismatches 3; Indels 2; Gaps 1;

Qy 6 VLMAMNIISKEKEIKWIGLP 26

A;Reference number: A89758; MUID:21311952; PMID:11418146

A;Accession: B89941  
A;Status: preliminary  
A:Molecule type: DNA

Db 14 VFSLNLSRPERE--WEGMP 32  
RESULT 3  
T37447  
Probable 8.5K protein - vaccinia virus (strain Ankara)

C;Species: vaccinia virus  
A;Variety: strain Ankara  
C;Date: 21-Jan-2000 #sequence\_revision 21-Jan-2000 #text\_change 21-Jan-2000

C;Accession: T37447  
R;Antoine, G.; Schefflinger, F.; Falkner, F.G.; Dörner, P.  
submitted to the EMBL Data Library, March 1997

A;Description: The complete genomic sequence of the Modified Vaccinia Ankara (MVA) strain

A;Reference number: Z20877

A;Accession: T37447

A;Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA

A;Residues: 1-74 <ANT>

A;Cross-references: EMBL:U94848; PIDN: AAB96552.1

A;Experimental source: strain Ankara

C;Genetics:

A;Note: MVA179R

Query Match 27.6%; Score 42; DB 2; Length 74;  
Best Local Similarity 52.6%; Pred. No. 43;  
Matches 10; Conservative 2; Mismatches 7; Indels 0; Gaps 0;

Qy 2 DALNVLMAMNIIISKEKEIK 20

Db 6 DVTNVEDIMNEIDREKEEI 24

#### RESULT 4

H90332  
Hypothetical protein SS08760 [imported] - Sulfolobus solfataricus transposon ISCl491

C;Species: Sulfolobus solfataricus

C;Date: 24-May-2001 #sequence\_revision 24-May-2001 #text\_change 24-May-2001

C;Accession: H90332

R;She, Q.; Singh, R.K.; Confalonieri, P.; Zivanovic, Y.; Allard, G.; Awayez, M.J.; Chan-  
Jong, I.; Jeffries, A.C.; Kozera, C.J.; Medina, N.; Peng, X.; Thi-Ngoc, H.P.; Redder, P.

arrett, R.A.; Ragan, M.A.; Sensen, C.W.; Van der Oost, J.

submitted to GenBank, April 2001

A;Description: Sulfolobus solfataricus complete genome.

A;Reference number: A99139

A;Accession: H90332

A;Status: preliminary

A:Molecule type: DNA

A;Residues: 1-76 <KUR>

A;Cross-references: GB:AE006641; NID:gl3814963; PIDN:AAK41919.1; GSPDB:GN00155

C;Genetics:

A;Gene: SS08760

Query Match 27.6%; Score 42; DB 2; Length 76;  
Best Local Similarity 40.0%; Pred. No. 44;  
Matches 8; Conservative 7; Mismatches 5; Indels 0; Gaps 0;

Qy 2 DALNVLMAMNIIISKEKEIK 21

Db 17 DKLTIVKAEIVITQEKREVE 36

#### RESULT 5

B89941  
conserved hypothetical protein SA1423 [imported] - Staphylococcus aureus (strain N315)

C;Species: Staphylococcus aureus

C;Date: 10-May-2001 #sequence\_revision 10-May-2001 #text\_change 02-Aug-2002

C;Accession: B89941

R;Kuroda, M.; Ohca, T.; Uchiyama, I.; Baba, T.; Yuzawa, H.; Kobayashi, I.; Cui, L.; Oguc  
ma, A.; Mizutani-Uji, Y.; Kobayashi, N.; Sawano, T.; Inoue, R.; Kaito, C.; Sekimizu, K.;

C.; Shiba, T.; Hattori, M.; Ogasawara, N.; Hayashi, H.; Hiramatsu, K.

Lancet 357, 1225-1240, 2001

A;Title: Whole genome sequencing of methicillin-resistant Staphylococcus aureus.

Query Match 26.3%; Score 40; DB 2; Length 76;  
Best Local Similarity 40.0%; Pred. No. 88;  
Matches 8; Conservative 6; Mismatches 6; Indels 0; Gaps 0;

Qy 2 DALNVLMAMNIIISKEKEIK 21

Db 17 DKLTIVKAEIVITQEKREVE 36

#### RESULT 7

B90299  
Hypothetical protein SS07998 [imported] - Sulfolobus solfataricus transposon ISCl491

C;Species: Sulfolobus solfataricus

C;Date: 24-May-2001 #sequence\_revision 24-May-2001 #text\_change 24-May-2001

C;Accession: B90299

R;She, Q.; Singh, R.K.; Confalonieri, P.; Zivanovic, Y.; Allard, G.; Awayez, M.J.; Chan-  
Jong, I.; Jeffries, A.C.; Kozera, C.J.; Medina, N.; Peng, X.; Thi-Ngoc, H.P.; Redder, P.

arrett, R.A.; Ragan, M.A.; Sensen, C.W.; Van der Oost, J.

submitted to GenBank, April 2001

A;Description: Sulfolobus solfataricus complete genome.

A;Reference number: A99139

A;Accession: B90299

A;Status: preliminary

A:Molecule type: DNA

A;Residues: 1-76 <KUR>

A;Cross-references: GB:AE006641; NID:gl3814637; PIDN:AAK41649.1; GSPDB:GN00155

C;Genetics:

A;Gene: SS07998

Query Match 26.3%; Score 40; DB 2; Length 69;  
Best Local Similarity 40.9%; Pred. No. 79;  
Matches 9; Conservative 4; Mismatches 9; Indels 0; Gaps 0;

Qy 9 AMNIIISKEKEIKWGLPTNSA 30

Db 21 AAKIIENEDVRAIVLPEGSA 42

#### RESULT 6

S66650

protease inhibitor BWI-1 - common buckwheat

C;Species: Fagopyrum esculentum (common buckwheat)

C;Date: 15-Feb-1997 #sequence\_revision 13-Mar-1997 #text\_change 09-May-1997

C;Accession: S66650

R;Belozersky, M.A.; Dunaevsky, Y.E.; Musolyanov, A.X.; Egorov, T.A.

FEBS Lett. 371, 264-266, 1995

A;Title: Complete amino acid sequence of the protease inhibitor from buckwheat seeds.

A;Reference number: S66650; MUID:96013147; PMID:7556606

A;Accession: S66650

A;Status: preliminary

A:Molecule type: protein

A;Residues: 1-69 <BEL>

C;Superfamily: eglin C

Query Match 27.6%; Score 42; DB 2; Length 96;  
Best Local Similarity 40.0%; Pred. No. 58;  
Matches 10; Conservative 6; Mismatches 9; Indels 0; Gaps 0;

Qy 2 DALNVLMAMNIIISKEKEIKWGLP 26

Db 72 ELVQVIGSMIVIVRESKENKEIPL 96

A;Reference number: A89758; MUID:21311952; PMID:11418146

A;Accession: B89941

A;Status: preliminary

A:Molecule type: DNA

A;Residues: 1-96 <KUR>

A;Cross-references: GB:BA000018; PID:gl3701393; PIDN:BAB42687.1; GSPDB:GN00149

A;Experimental source: strain N315

C;Genetics:

A;Gene: SA1423

C;Superfamily: Methanococcus jannaschii conserved hypothetical protein MJ0652

RESULT 8  
 B90331  
 hypothetical protein SSO8725 [imported] - Sulfolobus solfataricus transposon ISC1491  
 C/Species: Sulfolobus solfataricus  
 C/Date: 24-May-2001 #sequence\_revision 24-May-2001 #text\_change 24-May-2001  
 C/Accession: B90331  
 R/She, Q.; Singh, R.K.; Confalonieri, P.; Zivanovic, Y.; Allard, G.; Awayez, M.J.; Chan-  
 Jong, I.; Jeffries, A.C.; Kozera, C.J.; Medina, N.; Peng, X.; Thi-Ngoc, H.P.; Redder, F.  
 arrett, R.A.; Ragan, M.A.; Sensen, C.W.; Van der Oost, J.  
 submitted to GenBank, April 2001  
 A/Description: Sulfolobus solfataricus complete genome.  
 A/Reference number: A99139  
 A/Accession: B90331  
 A/Status: preliminary  
 A/Molecule type: DNA  
 A/Residues: 1-76 <KUR>  
 A/Cross-references: GB:AE006641; NID:gl3814947; PIDN:AAK41905.1; GSPDB:GNO0155  
 C/Genetics:  
 A/Gene: SSO8725  
  
 Query Match 26.3%; Score 40; DB 2; Length 76;  
 Best Local Similarity 40.0%; Pred. No. 88;  
 Matches 8; Conservative 6; Mismatches 6; Indels 0; Gaps 0;  
  
 Qy 2 DALNVLMMNNIIISKEKEIK 21  
 |||||:|:|:|:|:  
 Db 17 DKLTVAKELVITQEKREVE 36  
  
 RESULT 9  
 B81102  
 conserved hypothetical protein NMB1264 [imported] - Neisseria meningitidis (strain MC58  
 C/Species: Neisseria meningitidis  
 C/Date: 31-Mar-2000 #sequence\_revision 31-Mar-2000 #text\_change 19-Jan-2001  
 C/Accession: B81102  
 R/Tettelin, H.; Saunders, N.J.; Heidelberg, J.; Jeffries, A.C.; Nelson, K.E.; Eisen, J.A.;  
 Hickey, E.K.; Haft, D.H.; Salzberg, S.L.; White, O.; Fleischmann, R.D.; Dougherty, B.A.;  
 Ri, H.; Qin, H.; Vamathevan, J.; Gill, J.; Scarlato, V.; Maignani, V.; Pizza, M.  
 Science 287, 1809-1815, 2000  
 A/Authors: Grandi, G.; Sun, L.; Smith, H.O.; Fraser, C.M.; Moxon, E.R.; Rappuoli, R.; Ve-  
 A/Title: Complete genome sequence of Neisseria meningitidis serogroup B strain MC58.  
 A/Reference number: A81000; MUID:20175755; PMID:10710307  
 A/Accession: B81102  
 A/Status: preliminary  
 A/Molecule type: DNA  
 A/Residues: 1-90 <TET>  
 A/Cross-references: GB:AE002475; GB:AE002098; NID:g7226502; PIDN:AAF41641.1; PID:g7226500  
 A/Experimental source: serogroup B, strain MC58  
 C/Genetics:  
 A/Gene: NMB1264  
  
 Query Match 26.3%; Score 40; DB 2; Length 90;  
 Best Local Similarity 29.2%; Pred. No. 1.1e+02;  
 Matches 7; Conservative 8; Mismatches 9; Indels 0; Gaps 0;  
  
 Qy 2 DALNVLMMNNIIISKEKEIKWIGL 25  
 ||:|:|:|:|:|:|:|:  
 Db 53 NATHILKSYPEFAETTYLNNWIG 76  
  
 RESULT 10  
 D69406  
 hypothetical protein AF1253 - Archaeoglobus fulgidus  
 C/Species: Archaeoglobus fulgidus  
 C/Date: 05-Dec-1997 #sequence\_revision 05-Dec-1997 #text\_change 22-Oct-1999  
 C/Accession: D69406  
 R/Klenk, H.P.; Clayton, R.A.; Tomb, J.F.; White, O.; Nelson, K.E.; Ketchum, K.A.; Dodson-  
 Fleischmann, R.D.; Quackenbush, J.; Lee, N.H.; Sutton, G.G.; Gill, S.; Kirkness, E.F.;  
 Glödeke, A.; Zhou, L.; Overbeek, R.; Gocayne, J.D.; Weidman, J.F.; McDonald, L.  
 Nature 390, 364-370, 1997  
 A/Authors: Uitterback, T.; Cotton, M.D.; Spriggs, T.; Artiach, P.; Kaine, B.P.; Sykes, S.  
 Smith, H.O.; Woese, C.R.; Venter, J.C.  
 submitted to GenBank, April 1997

A;Title: The complete genome sequence of the hyperthermophilic, sulfate-reducing archaeo  
A;Reference number: A69250; MUID:98049343; PMID:9389475  
A;Accession: D69406  
A;Status: preliminary; nucleic acid sequence not shown; translation not shown  
A:Molecule type: DNA  
A;Residues: 1-64 <KLE>  
A;Cross-references: GB:AEO01018; GB:AEO00782; NID:g2689341; PIDN:AAB90006.1; PID:g264934

Query Match      25.7%; Score 39; DB 2; Length 64;  
Best Local Similarity 50.0%; Pred. No. 1e+02;  
Matches 10; Conservative 4; Mismatches 6; Indels 0; Gaps 0;

Oy    4 LNVLMANNIIISKEKEIKWI 23  
     |: :|: |:: ||||| |  
Db    42 LDKIEAVEEITTKAKKEIKKI 61

RESULT 11

hypothetical protein spr0300 [imported] - Streptococcus pneumoniae (strain R6)  
C:Species: Streptococcus pneumoniae  
C>Date: 22-Oct-2001 #sequence\_revision 22-Oct-2001 #text\_change 22-Oct-2001  
C:Accession: D97909  
R:Hoskins, J.A.; Alborn Jr., W.; Arnold, J.; Blaszcak, L.; Burgett, S.; DeHoff, B.S.; E.  
e, R.; LeBlanc, D.J.; Lee, L.N.; Lefkowitz, E.J.; Lu, J.; Matsushima, P.; McAhren, S.; M.  
Y, P.; Sun, P.M.; Winkler, M.E.  
J. Bacteriol. 183, 5709-5717, 2001  
A:Authors: Yang, Y.; Young-Bellido, M.; Zhao, G.; Zook, C.; Baltz, R.H.; Jaskunas, S.R.;  
A>Title: Genome of the Bacterium Streptococcus pneumoniae Strain R6.  
A;Reference number: A97872; MUID:21429245; PMID:11544234  
A;Accession: D97909  
A;Status: preliminary  
A:Molecule type: DNA  
A;Residues: 1-80 <KUR>  
A;Cross-references: GB:AEO07317; PIDN:AAK99104.1; PID:g15457854; GSPDB:GN00174  
C:Genetics:  
A;Gene: spr0300

Query Match      25.7%; Score 39; DB 2; Length 80;  
Best Local Similarity 41.2%; Pred. No. 1.3e+02;  
Matches 7; Conservative 5; Mismatches 5; Indels 0; Gaps 0;

Oy    13 ISKEKEIKWIGLPTNS 29  
     :||||:::| |  
Db    1 MKKEKKQLRYPCLKAGS 17

RESULT 12

AG1511  
hypothetical lin0631 [imported] - Listeria innocua (strain Clip11262)  
C:Species: Listeria innocua  
C>Date: 27-Nov-2001 #sequence\_revision 27-Nov-2001 #text\_change 27-Nov-2001  
C:Accession: AG1511  
R:Glaser, P.; Frangoul,  
.; Dominguez-Bernal, G.; Duchaud, E.; Durand, L.; Dussurget, O.; Entian, K.D.; Fsihi, H.  
D.; Jones, L.M.; Karst, U.  
Science 294, 849-852, 2001  
A:Authors: Kretz, J.; Kuhn, M.; Kunat, F.; Kurapkat, G.; Madueno, E.; Maitournam, A.; Ma  
ok, C.; Schlueter, T.; Simoes, N.; Tierrez, A.; Vazquez-Holand, J.A.; Voss, H.; Wehlend,  
A>Title: Comparative genomics of Listeria species.  
A;Reference number: AB1077; MUID:21537279; PMID:11679669  
A;Accession: AG1511  
A;Status: preliminary  
A:Molecule type: DNA  
A;Residues: 1-92 <GLA>  
A;Cross-references: GB:AL592022; PIDN:CAC95863.1; PID:g16413071; GSPDB:GN00178  
A;Experimental source: strain Clip11262  
C:Genetics:  
A;Gene: lin0631

Query Match      25.7%; Score 39; DB 2; Length 92;  
Best Local Similarity 32.0%; Pred. NO. 1.5e+02;  
Matches 8; Conservative 7; Mismatches 10; Indels 0; Gaps 0;

QY 4 LNVLMAMNIIISKEKEIKWIGLPTN 28  
| : : : : : : : : : : : : : : : :  
Db 49 LGVISTISVFFVKVKLVKILLGFN 73

## RESULT 13

hypoetical protein CP0930 [imported] - Chlamydothila pneumoniae (strain AR39)  
C:Species: Chlamydothila pneumoniae, Chlamydia pneumoniae  
C:Date: 31-Mar-2000 #sequence\_revision 31-Mar-2000 #text\_change 11-May-2000  
C:Accession: D81523  
R:Read, T.D.; Brunham, R.C.; Shen, C.; Gill, S.R.; Heidelberg, J.F.; White, O.; Hickey,  
C.; Dodson, R.; Gwin, M.; Nelson, W.; DeBoy, R.; Kolonay, J.; McClarty, G.; Salzberg,  
Nucleic Acids Res. 28, 1397-1406, 2000  
A:Title: Genome sequences of Chlamydia trachomatis MoPn and Chlamydia pneumoniae AR39.  
A:Reference number: A81500; MUID:20150255; PMID:10684935  
A:Accession: D81523  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-45 <REA>  
A:Cross-references: GB:AE002251; GB:AE002161; NID:g7189833; PIDN:AAF38713.1; PID:g718984  
A:Experimental source: strain AR39, HL cells  
C:Genetics:  
A:Gene: CP0930

Query Match 25.3%; Score 38.5; DB 2; Length 45;  
Best Local Similarity 43.5%; Pred. No. 81;  
Matches 10; Conservative 4; Mismatches 8; Indels 1; Gaps 1;

QY 1 YDALNVLAMNIIISKEKEIKW 23  
| : : : : : : : : : : : : : : : :  
Db 4 YEILRIIFMRF-FISFKEKEICYL 25

## RESULT 14

G64330  
chorismate mutase (EC 5.4.99.5) chain A - Methanococcus jannaschii  
C:Species: Methanococcus jannaschii  
C:Date: 10-Sep-1999 #sequence\_revision 10-Sep-1999 #text\_change 07-Mar-2003  
C:Accession: G64330  
R:Bult, C.J.; White, O.; Olsen, G.J.; Zhou, L.; Fleischmann, R.D.; Sutton, G.G.; Blake,  
Reich, C.I.; Overbeek, R.; Kirkness, E.F.; Weinstock, K.G.; Merrick, J.M.; Glodek, A.;  
Isaon, J.D.; Sadow, P.W.; Hanna, M.C.; Cotton, M.D.; Roberts, K.M.; Hurst, M.A.  
Science 273, 1058-1073, 1996  
A:Authors: Kaine, B.P.; Borodovsky, M.; Klenk, H.P.; Fraser, C.M.; Smith, H.O.; Woese,  
A:Title: Complete genome sequence of the methanogenic archaeon, Methanococcus jannaschii  
A:Reference number: A64300; MUID:96337999; PMID:8688087  
A:Accession: G64330  
A:Status: preliminary; nucleic acid sequence not shown; translation not shown  
A:Molecule type: DNA  
A:Residues: 1-99 <BUL>  
A:Cross-references: GB:U67480; GB:L77117; NID:g2826265; PIDN:AAB98234.1; PID:g1590979; T  
C:Genetics:  
A:Map position: REV233997-233698  
C:Superfamily: chorismate mutase of the AroQ class, prokaryotic type  
C:Keywords: intramolecular transferase; isomerase

Query Match 25.3%; Score 38.5; DB 1; Length 99;  
Best Local Similarity 34.6%; Pred. No. 2e+02;  
Matches 9; Conservative 7; Mismatches 9; Indels 1; Gaps 1;

QY 4 LNVLMAMNIIISKEKEIK-WIGLPTN 28  
| : : : : : : : : : : : : : : : :  
Db 20 LKLIARNSLAKDVAEIKNQLGIPIN 45

## RESULT 15

JQ1805  
Blir protein - vaccinia virus (strain WR)  
C:Species: vaccinia virus  
C:Date: 30-Sep-1993 #sequence\_revision 30-Sep-1993 #text\_change 08-Oct-1999  
C:Accession: JQ1805

R:Smith, G.L.; Chan, Y.S.; Howard, S.T.

J. Gen. Virol. 72, 1349-1376, 1991

A:Title: Nucleotide sequence of 42kbp of vaccinia virus strain WR from near the right in

A:Reference number: JQ1767; MUID:91259063; PMID:2045793

A:Accession: JQ1805

A:Molecule type: DNA

A:Residues: 1-72 <SMI>

A:Cross-references: DDBJ:D11079; NID:g222717; PIDN:BAA01841.1; PID:d1002317; PID:g222756

Query Match 25.0%; Score 38; DB 2; Length 72;  
Best Local Similarity 47.4%; Pred. No. 1.6e+02;  
Matches 9; Conservative 3; Mismatches 7; Indels 0; Gaps 0;

QY 2 DALNVLAMNIIISKEKEIK 20  
| : : : : : : : : : : : : : : : :  
Db 4 DVTNVEDIINEIDREKEIK 22

## RESULT 16

A42527  
Blir protein - vaccinia virus (strain Copenhagen)  
C:Species: vaccinia virus  
A:Note: host Homo sapiens (man)  
C:Date: 09-Nov-1990 #sequence\_revision 09-Nov-1990 #text\_change 08-Apr-1994  
C:Accession: A42527  
R:Johnson, G.P.  
submitted to GenBank, June 1990  
A:Reference number: A33172  
A:Accession: A42527  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-88 <JOH>

Query Match 25.0%; Score 38; DB 2; Length 88;  
Best Local Similarity 47.4%; Pred. No. 2e+02;  
Matches 9; Conservative 3; Mismatches 7; Indels 0; Gaps 0;

QY 2 DALNVLAMNIIISKEKEIK 20  
| : : : : : : : : : : : : : : : :  
Db 20 DVTNVEDIINEIDREKEIK 38

## RESULT 17

C64069  
virulence-associated protein vspD homolog H10450 - Haemophilus influenzae (strain Rd KW2)  
C:Species: Haemophilus influenzae  
C:Date: 10-Sep-1999 #sequence\_revision 10-Sep-1999 #text\_change 10-Sep-1999  
C:Accession: C64069  
R:Fleischmann, R.D.; Adams, M.D.; White, O.; Clayton, R.A.; Kirkness, E.F.; Kerlavage, A.  
; Gocayne, J.D.; Scott, J.; Shirley, R.; Liu, L.I.; Glodek, A.; Kelley, J.M.; Weidman, J.  
; D.M.; Brandon, R.C.; Fine, L.D.; Fritchman, J.L.; Fuhrmann, J.L.; Geoghegan, N.S.M.  
Science 269, 496-512, 1995  
A:Authors: Gnehm, C.L.; McDonald, L.A.; Small, K.V.; Fraser, C.M.; Smith, H.O.; Venter, J.  
A:Title: Whole-genome random sequencing and assembly of Haemophilus influenzae Rd.  
A:Reference number: A64000; MUID:95350630; PMID:7542800  
A:Accession: C64069  
A:Status: nucleic acid sequence not shown; translation not shown  
A:Molecule type: DNA  
A:Residues: 1-91 <TIGR>  
A:Cross-references: GB:U32728; GB:L42023; NID:g1573425; PIDN:AAC22108.1; PID:g1573426; T  
C:Superfamily: virulence-associated protein vspD

Query Match 25.0%; Score 38; DB 1; Length 91;  
Best Local Similarity 36.4%; Pred. No. 2.1e+02;  
Matches 8; Conservative 4; Mismatches 6; Indels 4; Gaps 1;

QY 2 DALNVLAMNIIISKEKEIKW 23  
| : : : : : : : : : : : : : : : :  
Db 50 DMANLFOAMNAL----KOLAWI 67

## RESULT 18

AF1152

A;Experimental source: strain DPC3147  
C;Genetics:  
A;Genome: plasmid pMRC01  
A;Note: ORF00054

Query Match 24.7%; Score 37.5; DB 2; Length 95;  
Best Local Similarity 57.9%; Pred. No. 2.6e+02;  
Matches 11; Conservative 0; Mismatches 7; Indels 1; Gaps 1;

Qy 9 AMNIISKEKEIKWI-GLP 26  
Db 14 AONTFEKEKKEDKIIQALP 32

RESULT 21  
T47420  
hypothetical protein T28A8.130 - Arabidopsis thaliana  
C;Species: Arabidopsis thaliana (mouse-ear cress)  
C;Date: 20-Apr-2000 #sequence\_revision 20-Apr-2000 #text\_change 20-Apr-2000  
C;Accession: T47420  
R;Purnelle, B.; Boutry, M.; Goffeau, A.; Mewes, H.W.; Rudd, S.; Lemcke, K.; Mayer, K.F.X  
submitted to the Protein Sequence Database, March 2000  
A;Reference number: Z24466  
A;Accession: T47420  
A;Status: preliminary  
A;Molecule type: DNA  
A;Residues: 1-84 <PUR>  
A;Cross-references: EMBL:AL162691  
A;Experimental source: cultivar Columbia; BAC clone T28A8  
C;Genetics:  
A;Map position: 3  
A;Introns: 32/2; 50/3  
A;Note: T28A8.130

Query Match 24.3%; Score 37; DB 2; Length 84;  
Best Local Similarity 28.6%; Pred. No. 2.7e+02;  
Matches 6; Conservative 8; Mismatches 7; Indels 0; Gaps 0;

Qy 6 VLAMNIISKEKEIKWIGLP 26  
Db 5 MMVMVAVSETNCVVKFTLP 25

RESULT 22  
B41869  
pro-sigma K processing regulatory protein bofa - Bacillus subtilis  
N;Alternate names: bypass-of-forespore A; pro-sigma-K processing machinery inhibitor bofA  
C;Species: Bacillus subtilis  
C;Date: 03-May-1994 #sequence\_revision 27-Jan-1995 #text\_change 15-Oct-1999  
C;Accession: B41869; S13789; S66053; H69595  
R;Ricca, E.; Cutting, S.; Losick, R.  
J. Bacteriol. 174, 3177-3184, 1992  
A;Title: Characterization of bofA, a gene involved in intercompartmental regulation of P1  
A;Reference number: A41869; MUID:92250411; PMID:1577688  
A;Accession: B41869  
A;Molecule type: DNA  
A;Residues: 1-54, 'G', 56-87 <RIC>  
R;Alonso, J.C.; Shirahige, K.; Ogasawara, N.  
Nucleic Acids Res. 18, 6771-6777, 1990  
A;Title: Molecular cloning, genetic characterization and DNA sequence analysis of the rec  
A;Reference number: S13786; MUID:91088245; PMID:2124672  
A;Accession: S13789  
A;Molecule type: DNA  
A;Residues: 1-87 <ALO>  
A;Cross-references: EMBL:X17014; NID:94533238; PIDN:CAA34880.1; PID:g40075  
R;Ogasawara, N.; Nakai, S.; Yoshikawa, H.  
DNA Res. 1, 1-14, 1994  
A;Title: Systematic sequencing of the 180 kilobase region of the Bacillus subtilis chrom  
A;Reference number: S65967; MUID:96051385; PMID:7584024  
A;Accession: S66053  
A;Status: nucleic acid sequence not shown; translation not shown  
A;Molecule type: DNA  
A;Residues: 1-87 <OGA>

C,Accession: E83815  
R,Takami, H.; Nakaone, K.; Takaki, Y.; Maeno, G.; Sasaki, R.; Masui, N.; Fujii, F.; Hirano,  
Nucleic Acids Res. 28, 4317-4331, 2000  
A,Title: Complete genome sequence of the alkaliphilic bacterium *Bacillus halodurans* and  
A,Reference number: A83650; MUID:20512582; PMID:11058132  
A,Accession: E83815  
A,Status: preliminary  
A,Molecule type: DNA  
A,Residues: 1-96 <STO>  
A,Cross-references: GB:AP001511; GB:BA000004; NID:gl0173727; PIDN:BAB05044.1; GSPDB:GN001  
A,Experimental source: strain C-125  
C,Genetics:  
C,Gene: BH1325  
C,Synfamily: Methanococcus jannaschii conserved hypothetical protein MJ0652

```

Query Match      24.3%; Score 37; DB 2; Length 96;
Best Local Similarity 36.0%; Pred. No. 3.1e+02;
Matches 9; Conservative 6; Mismatches 10; Indels 0; Gaps 0;

QY      2 DALNVLMMNNISKEKEIKWIGLP 26
      : : : : : : : : : :
Db      72 ELVQVIGNVILVYKESKEQKIETLP 96

RESULT 25
T45696
hypothetical protein F18L15.110 - Arabidopsis thaliana
C;Species: Arabidopsis thaliana (mouse-ear cress)
C;Date: 04-Feb-2000 #sequence_revision 04-Feb-2000 #text_change 04-Feb-2000
R;Accession: T45696
R;Choiene, N.; Robert, C.; Brottier, P.; Wincker, P.; Cattolico, L.; Artiguenave, F.; Sa
submitted to the Protein Sequence Database, November 1999
A;Reference number: Z23010

```

K;Cnolishé, N.; Robert, C.; Chonka, J.  
 submitted to the Protein Sequence Database, November 1999  
 A;Accession number: Z23010  
 A;Reference number: 223010  
 A;Accession: T45696  
 A;Status: preliminary  
 A;Molecule type: DNA  
 A;Residues: 1-69 <CHO>  
 A;Cross-references: EMBL:AL13298  
 A;Experimental source: cultivar Columbia; BAC clone F18L15  
 C;Genetics:  
 A;Map position: 3  
 A;Introns: 40/3  
 A;Note: F18L15.110

```

A:Experimental Source: Curatorial Committee;
C:Genetics:
A:Map position: 3
A:Introns: 40/3
A:Note: F18L15.110

Query Match          24.0%; Score 36.5; DB 2; Length 69;
Best Local Similarity 38.5%; Pred. NO. 2.6e+02;
Matches 10; Conservative 3; Mismatches 6; Indels 7; Gaps 1;

QY      7  LMANNILSKERKE-----IKWIGL 25
      : : : : :
Db      28  IVAKNMAKKEKIEQVISMGIMWGL 53

RESULT 26
T07291
hypothetical protein 42c - Chlorella vulgaris chloroplast
C:Species: Chloroplast Chlorella vulgaris
C:Date: 14-May-1999 #sequence_revision 14-May-1999 #text_change 21-Jul-2000
C:Accession: T07291

```

A:Accession: T07291  
 A:Status: preliminary; translated from GB/EMBL/DBJ  
 A:Molecule type: DNA  
 A:Residues: 1-42 <WAK>  
 A:Cross-references: EMBL:AB01684; NID:G2224352; PIDN:BAAS7939.1; PID:G2224455  
 C:Genetics:  
 A:Genome: chloroplast  
 C:Keywords: chloroplast  
 Query Match 23.7%; Score 36; DB 2; Length 42;

Best Local Similarity 47.1%; Pred. No. 1.8e+02; Mismatches 3; Gaps 0; Indels 0;

Qy 7 LMAWNIISKEKEIKWI 23  
| : ||:||||| :  
Db 16 LFLIFFSKKKKEIKQV 32

RESULT 27  
A89995  
delta-hemolysin [imported] - Staphylococcus aureus (strain N315)  
C:Species: Staphylococcus aureus  
C:Date: 10-May-2001 #sequence\_revision 10-May-2001 #text\_change 22-Oct-2001  
C:Accession: A89995  
R:Kuroda, M.; Ohta, T.; Uchiyama, I.; Baba, T.; Yuzawa, H.; Kobayashi, I.; Cui, L.; Oguc  
ma, A.; Mizutani-Ui, Y.; Kobayashi, N.; Sawano, T.; Inoue, R.; Kaito, C.; Sekimizu, K.;  
C.; Shiba, T.; Hattori, M.; Ogasawara, N.; Hayashi, H.; Hiramatsu, K.  
Lancet 357, 1225-1249, 2001  
A:Title: Whole genome sequencing of methicillin-resistant Staphylococcus aureus.  
A:Reference number: A89758; MUID:21311952; PMID:11418146  
A:Accession: A89995  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-44 <KUR>  
A:Cross-references: GB:BA000018; PID:g13701829; PIDN:BA843122.1; GSPDB:GN00149  
A:Experimental source: strain N315  
C:Genetics:  
A:Gene: hld

Query Match 23.7%; Score 36; DB 2; Length 44;  
Best Local Similarity 44.4%; Pred. No. 1.8e+02; Mismatches 3; Gaps 0; Indels 0;  
Matches 8; Conservative 3; Mismatches 7; Indels 0;

Qy 6 VLMANNIISKEKEIKWI 23  
: || : ||| : |||  
Db 17 ISMAQDIISIGDLVKWI 34

RESULT 28  
B64368  
hypothetical protein MJ0546 - Methanococcus jannaschii  
C:Species: Methanococcus jannaschii  
C:Date: 13-Sep-1996 #sequence\_revision 13-Sep-1996 #text\_change 21-Jul-2000  
C:Accession: B64368  
R:Bult, C.J.; White, O.; Olsen, G.J.; Zhou, L.; Fleischmann, R.D.; Sutton, G.G.; Blake,  
; Reich, C.I.; Overbeek, R.; Kirkness, E.F.; Weinstock, K.G.; Merrick, J.M.; Glodek, A.;  
rson, J.D.; Sadow, P.W.; Hanna, M.C.; Cotton, M.D.; Roberts, K.M.; Hurst, M.A.  
Science 273, 1058-1073, 1996  
A:Authors: Kaine, B.P.; Borodovsky, M.; Klenk, H.P.; Fraser, C.M.; Smith, H.O.; Woese, C  
A:Title: Complete genome sequence of the methanogenic archaeon, Methanococcus jannaschii  
A:Reference number: A64300; MUID:96337999; PMID:8688087  
A:Accession: B64368  
A:Status: preliminary; nucleic acid sequence not shown; translation not shown  
A:Molecule type: DNA  
A:Residues: 1-56 <BUU>  
A:Cross-references: GB:U67504; GB:L77117; NID:g1591248; PIDN:AA898548.1; PID:g1591251; T  
C:Genetics:  
A:Map position: REV483625-483455  
C:Superfamily: hypothetical protein MJ0546

Query Match 23.7%; Score 36; DB 2; Length 56;  
Best Local Similarity 35.0%; Pred. No. 2.4e+02; Mismatches 7; Gaps 1;  
Matches 7; Conservative 7; Mismatches 4; Indels 2; Gaps 1;

Qy 4 LNVLMANNII--SKEKEIK 21  
||::: : || : ||  
Db 13 LNIILGKIVMLQKELEVK 32

RESULT 29  
JC1386  
homeotic protein EgHbx1 - tapeworm (Echinococcus granulosus) (fragment)  
C:Species: Echinococcus granulosus

C:Date: 30-Sep-1993 #sequence\_revision 30-Sep-1993 #text\_change 17-Oct-1997  
C:Accession: JC1386  
R:Oliver, G.; Vispo, M.; Mailhos, A.; Martinez, C.; Sosa-Pineda, B.; Fielitz, W.; Ehrlich  
Gene 121, 337-342, 1992  
A:Title: Homeoboxes in flatworms.  
A:Reference number: JC1386; MUID:93077050; PMID:1359988  
A:Accession: JC1386  
A:Molecule type: DNA  
A:Residues: 1-60 <OLI>  
A:Cross-references: EMBL:X66817  
C:Superfamily: unassigned homeobox proteins; homeobox homology  
C:Keywords: DNA binding; homeobox; nucleus; transcription regulation  
F:2-58/Domain: homeobox homology <HOX>

Query Match 23.7%; Score 36; DB 2; Length 60;  
Best Local Similarity 25.0%; Pred. No. 2.6e+02; Mismatches 7; Conservative 7; Mismatches 8; Indels 6; Gaps 1;  
Matches 7; Conservative 7; Mismatches 8; Indels 6; Gaps 1;

Qy 1 YDNLVLMANNIISKEKK-----EIKW 22  
| : ||: ||| : ||  
Db 29 YERLNALSLNLTETQVQWIFQNRRTKW 56

RESULT 30  
AE3378  
hypothetical protein BMEI1011 [imported] - Brucella melitensis (strain 16M)  
C:Species: Brucella melitensis  
C:Date: 01-Feb-2002 #sequence\_revision 01-Feb-2002 #text\_change 01-Feb-2002  
C:Accession: AE3378  
R:DelVecchio, V.G.; Kapral, V.; Redkar, R.J.; Patra, G.; Mujer, C.; Los, T.; Ivanova, P.;  
; Mazur, M.; Gotsman, E.; Selkov, E.; Elzer, P.H.; Hagius, S.; O'Callaghan, D.; Letessier  
Proc. Natl. Acad. Sci. U.S.A. 99, 443-448, 2002  
A:Title: The genome sequence of the facultative intracellular pathogen Brucella melitensis  
A:Reference number: AD3252; PMID:11756688  
A:Accession: AE3378  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-67 <KUR>  
A:Cross-references: GB:AE008917; PIDN:AAL52192.1; PID:g17982973; GSPDB:GN00190  
A:Experimental source: strain 16M  
C:Genetics:  
A:Gene: BMEI1011  
A:Map position: 1

Query Match 23.7%; Score 36; DB 2; Length 67;  
Best Local Similarity 47.1%; Pred. No. 3e+02; Mismatches 8; Conservative 3; Mismatches 6; Indels 0; Gaps 0;  
Matches 8; Conservative 3; Mismatches 6; Indels 0; Gaps 0;

Qy 4 LNVLMANNIISKEKEI 20  
| : ||: ||| : ||  
Db 30 LKQLFSMGLISREPKTI 46

RESULT 31  
T48368  
hypothetical protein F12E4.200 - Arabidopsis thaliana  
C:Species: Arabidopsis thaliana (mouse-ear cress)  
C:Date: 20-Apr-2000 #sequence\_revision 20-Apr-2000 #text\_change 20-Apr-2000  
C:Accession: T48368  
R:Bevan, M.; Pohl, T.; Weizenegger, T.; Bancroft, I.; Mewes, H.W.; Rudd, S.; Lemcke, K.;  
submitted to the Protein Sequence Database, March 2000  
A:Reference number: Z24492  
A:Accession: T48368  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-73 <BEV>  
A:Cross-references: EMBL:AL162751  
A:Experimental source: cultivar Columbia; BAC clone F12E4  
C:Genetics:  
A:Map position: 5  
A:Introns: 30/1; 56/3  
A:Note: F12E4.200

```
Query Match      23.7%; Score 36; DB 2; Length 73;
Best Local Similarity 37.5%; Pred. No. 3.3e+02;
Matches 6; Conservative 5; Mismatches 5; Indels 0; Gaps 0;

QY 13 ISKEKEIKWIGLPTN 28
Db 32 IAWQEMRKWVGDFSN 47

RESULT 32
G69286
hypothetical protein AF0295 - Archaeoglobus fulgidus
C:Species: Archaeoglobus fulgidus
C>Date: 05-Dec-1997 #sequence_revision 05-Dec-1997 #text_change 22-Oct-1999
C:Accession: G69286
R;Klenk, H.P.; Clayton, R.A.; Tomb, J.F.; White, O.; Nelson, K.E.; Ketchum, K.A.; Dodson
.; Fleischmann, R.D.; Quackenbush, J.; Lee, N.H.; Sutton, G.G.; Gill, S.; Kirkness, E.F.
Glodek, A.; Zhou, L.; Overbeek, R.; Gocayne, J.D.; Weidman, J.F.; McDonald, L.
Nature 390, 364-370, 1997
A:Authors: Utterback, T.; Cotton, M.D.; Spriggs, T.; Artiach, P.; Kaine, B.P.; Sykes, S.
Smith, H.O.; Woese, C.R.; Venter, J.C.
A:Title: The complete genome sequence of the hyperthermophilic, sulfate-reducing archae
A:Reference number: A69250; MUID:98049343; PMID:9389475
A:Accession: G69286
A:Status: preliminary; nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-74 <KLE>
A:CROSS-references: GB:AE001084; GB:AE000782; NID:92699407; PIDN:AAB90939.1; PID:g265034

Query Match      23.7%; Score 36; DB 2; Length 74;
Best Local Similarity 47.4%; Pred. No. 3.3e+02;
Matches 9; Conservative 2; Mismatches 8; Indels 0; Gaps 0;

QY 4 LNVLMAMNIISKEKEIKW 22
Db 3 VQVLEAGKLIISPNEKVIW 21

RESULT 33
A24500
Real hypothetical protein 1 - Neisseria gonorrhoeae
C:Species: Neisseria gonorrhoeae
C>Date: 30-Jun-1988 #sequence_revision 30-Jun-1988 #text_change 23-Mar-1995
C:Accession: A24500
R;Correia, F.F.; Inouye, S.; Inouye, M.
J. Bacteriol. 167, 1009-1015, 1986
A:Title: A 26-base-pair repetitive sequence specific for Neisseria gonorrhoeae and Neis
A:Reference number: A91818; MUID:86304144; PMID:3091577
A:Accession: A24500
A:Molecule type: DNA
A:Residues: 1-75 <COR>
A:Experimental source: subclone pNG273

Query Match      23.7%; Score 36; DB 2; Length 75;
Best Local Similarity 29.2%; Pred. No. 3.4e+02;
Matches 7; Conservative 9; Mismatches 8; Indels 0; Gaps 0;

QY 4 LNVLMAMNIISKEKEIKWIGLPT 27
Db 49 LSLCLSDSSISRSKRAIRCLALST 72

RESULT 34
S33665
hypothetical protein 2 - Pseudomonas aeruginosa phage CTX
C:Species: Pseudomonas aeruginosa phage CTX
C>Date: 31-Dec-1993 #sequence_revision 31-Dec-1993 #text_change 08-Oct-1999
C:Accession: S33665
R;Hayashi, T.; Matsu moto, H.; Ohnishi, M.; Terawaki, Y.
Mol. Microbiol. 7, 657-667, 1993
A:Title: Molecular analysis of a cytotoxin-converting phage, phi-CTX, of Pseudomonas aer
A:Reference number: S33665; MUID:93225809; PMID:8469112
A:Accession: S33665
```

```
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-77 <HAY>
A:CROSS-references: EMBL:DI3409; NID:g217776; PIDN:BAA02674.1; PID:d1003179; PID:g2168161
A>Note: the authors translated the initiation codon GTG for residue 1 as Val
C:Genetics:
A:Start codon: GTG

Query Match      23.7%; Score 36; DB 2; Length 77;
Best Local Similarity 54.5%; Pred. No. 3.5e+02;
Matches 6; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY 20 IKWIGLPTNSA 30
Db 31 VPWVGLPTQLA 41

RESULT 35
JQ1781
Salp6R 8.8K protein - vaccinia virus (strain WR and Ankara)
C:Species: vaccinia virus
C>Date: 30-Sep-1993 #sequence_revision 30-Sep-1993 #text_change 18-Feb-2000
C:Accession: JQ1781; D40897; T37429
R;Smith, G.L.; Chan, Y.S.; Howard, S.T.
J. Gen. Virol. 72, 1349-1376, 1991
A:Title: Nucleotide sequence of 42kbp of vaccinia virus strain WR from near the right in
A:Reference number: JQ1767; MUID:91259063; PMID:2045793
A:Accession: JQ1781
A:Molecule type: DNA
A:Residues: 1-78 <SMI>
A:CROSS-references: DDBJ:DI1079; NID:g222717; PIDN:BAA01817.1; PID:d1002293; PID:g222732
R;Blasco, R.; Cole, N.B.; Mose, B.
J. Virol. 65, 4598-4608, 1991
A:Title: Sequence analysis, expression, and deletion of a vaccinia virus gene encoding a
A:Reference number: A40897; MUID:91332999; PMID:1870190
A:Accession: D40897
A:Molecule type: DNA
A:Residues: 1-78 <BLA>
A:CROSS-references: GB:M72474; NID:g335761; PIDN:AAA48310.1; PID:g335765
R;Antoine, G.; Scheiflinger, F.; Falkner, F.G.; Dörner, F.
Submitted to the EMBL Data Library, March 1997
A:Description: The complete genomic sequence of the Modified Vaccinia Ankara (MVA) strain
A:Reference number: Z20877
A:Accession: T37429
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-78 <ANT>
A:CROSS-references: EMBL:U94848; PIDN:AAB96536.1
A:Experimental source: strain Ankara
C:Genetics:
A>Note: MVA156R

Query Match      23.7%; Score 36; DB 2; Length 78;
Best Local Similarity 35.3%; Pred. No. 3.5e+02;
Matches 6; Conservative 7; Mismatches 4; Indels 0; Gaps 0;

QY 5 NVLMAMNIISKEKEIK 21
Db 42 DVLLAQSVAVEAKDKV 58

RESULT 36
AC1217
hypothetical protein lm0139 [imported] - Listeria monocytogenes (strain EGD-e)
C:Species: Listeria monocytogenes
C>Date: 27-Nov-2001 #sequence_revision 27-Nov-2001 #text_change 27-Nov-2001
C:Accession: AC1217
R;Glaser, P.; Frangeul, L.; Buchrieser, C.; Amend, A.; Baquero, F.; Berche, P.; Bloecker
.; Dominguez-Bernal, G.; Duchaud, E.; Durand, L.; Dussurget, O.; Entian, K.D.; Fsihi, H.
D.; Jones, L.M.; Karst, U.
Science 294, 849-852, 2001
A:Authors: Kreft, J.; Kuhn, M.; Kunst, F.; Kurapkat, G.; Madueno, E.; Maitournam, A.; Ma
Ok, C.; Schlueter, T.; Simoes, N.; Tierrez, A.; Vazquez-Boland, J.A.; Voss, H.; Wehland,
```



A;Title: Comparative genomics of *Listeria* species.  
 A;Reference number: AB1077; MUID:21537279; PMID:11679669  
 A;Accession: AC1217  
 A;Status: preliminary  
 A;Molecule type: DNA  
 A;Residues: 1-79 <GLA>  
 A;Cross-references: GB:NC\_003210; PIDN:CAC99217.1; PID:g16410555; GSPDB:GN00177  
 A;Experimental source: strain EGD-e  
 C;Genetics:  
 A;Gene: lmo1139

Query Match 23.7%; Score 36; DB 2; Length 79;  
 Best Local Similarity 43.8%; Pred. No. 3.6e+02;  
 Matches 7; Conservative 3; Mismatches 6; Indels 0; Gaps 0;

QY 10 MNIIISKEKKEIKWGL 25  
 :|: ||| ||| :|:  
 Db 33 VNLSYSEKDEQKWLQL 48

RESULT 37  
 AD1134  
 hypothetical protein lmo0475 [imported] - *Listeria* monocytogenes (strain EGD-e)  
 C;Species: *Listeria* monocytogenes  
 C;Date: 27-Nov-2001 #sequence\_revision 27-Nov-2001 #text\_change 27-Nov-2001  
 C;Accession: AD1134  
 R;Glaser, P.; Frangeul, L.; Buchrieser, C.; Amend, A.; Baquero, F.; Berche, P.; Bloecker, D.; Dominguez-Bernal, G.; Duchaud, E.; Durand, L.; Dussurget, O.; Entian, K.D.; Fsihi, H.; Jones, L.M.; Karst, U.; Science 294, 849-852, 2001  
 A;Authors: Kreft, J.; Kuhn, M.; Kunst, F.; Kurapkat, G.; Madueno, E.; Maitournam, A.; Mak, C.; Schluteter, T.; Simoes, N.; Tierrez, A.; Vazquez-Boland, J.A.; Voss, H.; Wehlend, A.; Title: Comparative Genomics of *Listeria* species.  
 A;Reference number: AB1077; MUID:21537279; PMID:11679669  
 A;Accession: AD1134  
 A;Status: preliminary  
 A;Molecule type: DNA  
 A;Residues: 1-80 <GLA>  
 A;Cross-references: GB:NC\_003210; PIDN:CAC98554.1; PID:g16409851; GSPDB:GN00177  
 A;Experimental source: strain EGD-e  
 C;Genetics:  
 A;Gene: lmo0475

Query Match 23.7%; Score 36; DB 2; Length 80;  
 Best Local Similarity 40.0%; Pred. No. 3.6e+02;  
 Matches 6; Conservative 5; Mismatches 4; Indels 0; Gaps 0;

QY 8 MANNIISKEKKEIKW 22  
 :|: ||| :|:  
 Db 46 LATSVNVKKKEVILW 60

RESULT 38  
 B64888  
 ynfA protein - *Escherichia coli* (strain K-12)  
 C;Species: *Escherichia coli*  
 C;Date: 12-Sep-1997 #sequence\_revision 17-Sep-1997 #text\_change 01-Mar-2002  
 C;Accession: B64888  
 R;Blattner, F.R.; Plunkett III, G.; Bloch, C.A.; Perna, N.T.; Burland, V.; Riley, M.; Cohen, A.; Rose, D.J.; Mau, B.; Shao, Y.; Science 277, 1453-1462, 1997  
 A;Title: The complete genome sequence of *Escherichia coli* K-12.  
 A;Reference number: A64720; MUID:97426617; PMID:9278503  
 A;Accession: B64888  
 A;Status: nucleic acid sequence not shown; translation not shown  
 A;Molecule type: DNA  
 A;Residues: 1-88 <BLAT>  
 A;Cross-references: GB:AE000234; GB:U00096; NID:g1787633; PIDN:AAC74457.1; PID:g1787639;  
 A;Experimental source: strain K-12, substrain MG1655  
 C;Genetics:  
 A;Gene: ynfA

Query Match 23.7%; Score 36; DB 2; Length 88;

Best Local Similarity 44.4%; Pred. No. 4e+02;  
 Matches 8; Conservative 2; Mismatches 6; Indels 2; Gaps 1;  
 QY 5 NVLMAMNIIISKEKKEIKW 22  
 :|: ||| ||| :|:  
 Db 5 NIL--RNIFPKSKDTLKW 20

RESULT 39  
 C64909  
 ydfK protein - *Escherichia coli* (strain K-12)  
 C;Species: *Escherichia coli*  
 C;Date: 12-Sep-1997 #sequence\_revision 17-Sep-1997 #text\_change 01-Mar-2002  
 C;Accession: C64909  
 R;Blattner, F.R.; Plunkett III, G.; Bloch, C.A.; Perna, N.T.; Burland, V.; Riley, M.; Cohen, A.; Rose, D.J.; Mau, B.; Shao, Y.; Science 277, 1453-1462, 1997  
 A;Title: The complete genome sequence of *Escherichia coli* K-12.  
 A;Reference number: A64720; MUID:97426617; PMID:9278503  
 A;Accession: C64909  
 A;Status: nucleic acid sequence not shown; translation not shown  
 A;Molecule type: DNA  
 A;Residues: 1-88 <BLAT>  
 A;Cross-references: GB:AE000252; GB:U00096; NID:g2367119; PIDN:AAC74617.1; PID:g1787826;  
 A;Experimental source: strain K-12, substrain MG1655

Query Match 23.7%; Score 36; DB 2; Length 88;  
 Best Local Similarity 44.4%; Pred. No. 4e+02;  
 Matches 8; Conservative 2; Mismatches 6; Indels 2; Gaps 1;

QY 5 NVLMAMNIIISKEKKEIKW 22  
 :|: ||| ||| :|:  
 Db 5 NIL--RNIFPKSKDTLKW 20

RESULT 40  
 E69951  
 dihydrodipicolinate reductase homolog ygeI - *Bacillus subtilis*  
 C;Species: *Bacillus subtilis*  
 C;Date: 05-Dec-1997 #sequence\_revision 05-Dec-1997 #text\_change 20-Jun-2000  
 C;Accession: E69951  
 R;Kunst, F.; Ogawara, N.; Moszer, I.; Albertini, A.M.; Alloni, G.; Azevedo, V.; Bertero, C.; Bron, S.; Brouillet, S.; Bruschi, C.V.; Caldwell, B.; Capuano, V.; Carter, N.M.; Choi, A.; Ehrlich, S.D.; Emmerson, P.T.; Entian, K.D.; Errington, J.; Fabret, C.; Ferrari, E.; Nature 390, 249-256, 1997  
 A;Authors: Foulger, D.; Fritz, C.; Fujita, M.; Fujita, Y.; Fuma, S.; Galizzi, A.; Galler, J.; Harwood, C.R.; Henaut, A.; Hilbert, H.; Holsappel, S.; Hosono, S.; Hullo, M.F.; Koetter, P.; Koningsstein, G.; Krogh, S.; Kumano, M.; Kurita, K.; Lapidus, A.; Lardinois, A.; Authors: Lauber, J.; Lazarevic, V.; Lee, S.M.; Levine, A.; Liu, H.; Masuda, S.; Maueel, Y.; M.; Ogawa, K.; Ogiwara, A.; Oudega, B.; Park, S.H.; Parro, V.; Pohl, T.M.; Portetelle, Rieger, M.; Rivolta, C.; Rocha, E.; Roche, B.; Rose, M.; Sadaie, Y.; Sato, T.; Scanlon, A.; Authors: Schleich, S.; Schroeter, R.; Scoffone, F.; Segiguchi, J.; Sekowaka, A.; Seror, akeuchi, M.; Tamakoshi, A.; Tanaka, T.; Terpstra, P.; Tognoni, A.; Tosato, V.; Uchiyama, T.; Winters, P.; Wipat, A.; Yamamoto, H.; Yamane, K.; Yasumoto, K.; Yata, K.; Yoshida, K.; Authors: Yoshikawa, H.F.; Zumstein, E.; Yoshikawa, H.; Danchin, A.  
 A;Title: The complete genome sequence of the Gram-positive bacterium *Bacillus subtilis*.  
 A;Reference number: A69580; MUID:98044033; PMID:9384377  
 A;Accession: E69951  
 A;Status: preliminary; nucleic acid sequence not shown; translation not shown  
 A;Molecule type: DNA  
 A;Residues: 1-96 <KUN>  
 A;Cross-references: GB:Z99117; GB:AL009126; NID:g2634966; PIDN:CAB14507.1; PID:g2635011  
 A;Experimental source: strain 168  
 C;Genetics:  
 A;Gene: ygeI  
 C;Superfamily: Methanococcus jannaschii conserved hypothetical protein MJ0652

Query Match 23.7%; Score 36; DB 2; Length 96;  
 Best Local Similarity 53.3%; Pred. No. 4.4e+02;  
 Matches 8; Conservative 2; Mismatches 5; Indels 0; Gaps 0;

QY 12 IISKEKKEIKWGLP 26  
 :|: ||| ||| :|:

F85981  
hypothetical protein yhbY [imported] - Escherichia coli (strain O157:H7, substrain EDL93)  
C/Species: Escherichia coli  
C/Date: 16-Feb-2001 #sequence\_revision 16-Feb-2001 #text\_change 17-May-2002  
C/Accession: F85981  
R/Perna, N.T.; Plunkett III, G.; Burland, V.; Mau, B.; Glasner, J.D.; Rose, D.J.; Mayhew  
iller, L.; Grobbeck, E.J.; Davis, N.W.; Lam, A.; Dimalanta, E.; Potamousis, K.; Apodaca,  
Nature 409, 528-533, 2001  
nt. Genome sequence of enterohemorrhagic Escherichia coli O157:H7.

A:Titl

A;Reference number: A96039; MUID:21368234; PMID:11474104

A;Contents: annotation

C;Genetics:

A;Gene: SMB21004

A;Genome: plasmid

Query Match 23.7%; Score 36; DB 2; Length 98;  
Best Local Similarity 55.6%; Pred. No. 4.5e+02;  
Matches 5; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

Qy 18 KEIKWIGLP 26

Db :||| |.

3 EEVNRGVP 11

Search completed: February 17, 2004, 10:58:18

Job time : 19.1188 secs

THIS PAGE BLANK (user)

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: February 17, 2004, 10:50:13 ; Search time 4.22772 Seconds  
(without alignments)  
159.230 Million cell updates/sec

Title: US-09-900-147-7

Perfect score: 31

Sequence: 1 ALNVLMA 7

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283308 seqs, 96168682 residues

Total number of hits satisfying chosen parameters: 37673

Minimum DB seq length: 0

Maximum DB seq length: 100

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

1: pir1:\*\*

2: pir2:\*\*

3: pir3:\*\*

4: pir4:\*\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	25	80.6	53	2	D83832
2	25	80.6	98	2	NADH2 dehydrogenas
3	24	77.4	50	2	ccoh protein - Par
4	24	77.4	94	2	yaIE protein - Esc
5	24	77.4	94	2	yaIE protein [simi
6	24	77.4	94	2	yaIE protein [simi
7	24	77.4	98	2	NADH2 dehydrogenas
8	24	77.4	98	2	NADH2 dehydrogenas
9	23	74.2	51	2	hypothetical prote
10	23	74.2	96	1	cytochrome c553 pr
11	23	74.2	96	2	hypothetical prote
12	22	71.0	54	2	hemolysin - Proteu
13	22	71.0	55	2	proline/betaine tr
14	22	71.0	56	2	protein kinase AK1
15	22	71.0	63	2	MHC class II histo
16	22	71.0	65	2	probable membrane
17	22	71.0	68	2	hypothetical prote
18	22	71.0	69	2	juvenile-hormone e
19	22	71.0	77	2	ORF X protein - hu
20	22	71.0	82	2	hypothetical prote
21	22	71.0	85	2	oxidoreductase hom
22	22	71.0	87	2	ycwE protein - Bac
23	22	71.0	89	2	NADH2 dehydrogenas
24	22	71.0	95	2	hypothetical prote
25	22	71.0	98	1	OXBO4L
26	22	71.0	98	2	NADH2 dehydrogenas
27	22	71.0	98	2	NADH2 dehydrogenas
28	22	71.0	98	2	elongin C - fruit
29	21	67.7	36	2	hypothetical prote

30 21 67.7 57 2 PN0624  
31 21 67.7 57 2 T16930  
32 21 67.7 61 2 A69337  
33 21 67.7 63 2 H90118  
34 21 67.7 63 2 I38244  
35 21 67.7 65 2 H90804  
36 21 67.7 65 2 E91196  
37 21 67.7 65 2 D86043  
38 21 67.7 65 2 E85613  
39 21 67.7 66 2 H98064  
40 21 67.7 71 2 AH0390  
41 21 67.7 72 2 AC2335  
42 21 67.7 75 2 B75106  
43 21 67.7 75 2 D71028  
44 21 67.7 78 2 S52172  
45 21 67.7 80 2 E85664

#### ALIGNMENTS

##### RESULT 1

D83832

hypothetical protein BH1460 [imported] - Bacillus halodurans (strain C-125)

C;Species: Bacillus halodurans

C;Date: 01-Dec-2000 #sequence\_revision 01-Dec-2000 #text\_change 15-Jun-2001

C;Accession: D83832

R;Takami, H.; Nakasone, K.; Takaki, Y.; Maeno, G.; Sasaki, R.; Masui, N.; Fuji, F.; Hiran

Nucleic Acids Res. 28, 4317-4331, 2000

A;Title: Complete genome sequence of the alkaliphilic bacterium Bacillus halodurans and

A;Reference number: A83650; MUID:20512582; PMID:11058132

A;Accession: D83832

A;Status: preliminary

A;Molecule type: DNA

A;Residues: 1-53 <STO>

A;Cross-references: GB:AP001512; GB:BA000004; NID:gl0174030; PIDN:BA05179.1; GSPDB:GN001

A;Experimental source: strain C-125

C;Genetics:

A;Gene: BH1460

Query Match 80.6%; Score 25; DB 2; Length 53;

Best Local Similarity 85.7%; Pred. No. 29;

Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 ALNVLMA 7

Db 15 ALNVLAA 21

##### RESULT 2

S41828

NADH2 dehydrogenase (ubiquinone) (EC 1.6.5.3) chain 4L - blue whale mitochondrion

N;Alternate names: NADH-ubiquinone oxidoreductase chain 4L

C;Species: mitochondrion Balaeoptera musculus (blue whale)

C;Date: 07-Oct-1994 #sequence\_revision 07-Oct-1994 #text\_change 03-Jun-2002

C;Accession: S41828

R;Arnason, U.; Gullberg, A.

J. Mol. Evol. 37, 312-322, 1993

A;Title: Comparison between the complete mtDNA sequences of the blue and the fin whale,

A;Reference number: S41820; MUID:94141932; PMID:8308901

A;Accession: S41828

A;Molecule type: DNA

A;Residues: 1-98 <ARN>

A;Cross-references: EMBL:X72204; NID:g414126; PIDN:CAA51003.1; PID:g575317

C;Genetics:

A;Gene: NADH4L

A;Genome: mitochondrion

A;Genetic code: SGC1

A;Start codon: GTG

C;Superfamily: NADH dehydrogenase (ubiquinone) chain 4L

C;Keywords: membrane-associated complex; mitochondrion; NAD; oxidative phosphorylation;

Query Match 80.6%; Score 25; DB 2; Length 98;

Tue Feb 17 11:55:56 2004

us-09-900-147-7.rpr

```
Best Local Similarity 83.3%; Pred. No. 57;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 2 LNVLM 7
Db 6 MNVLM 11

RESULT 3
cchH protein - Paracoccus denitrificans (fragment)
C:Species: Paracoccus denitrificans
C>Date: 24-Oct-1998 #sequence_revision 24-Oct-1998 #text_change 21-Jul-2000
C:Accession: S77603
R:de Gier, J.W.; Schepper, M.; Reijnders, W.N.M.; van Dyck, S.J.; Slotboom, D.J.; Warne,
Mol. Microbiol. 20, 1247-1260, 1996
A:Title: Structural and functional analysis of aa(3)-type and cbb(3)-type cytochrome c
A:Reference number: S77595; MUID:96405647; PMID:8809776
A:Accession: S77603
A:Status: nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-50 <DEA>
A:Cross-references: EMBL:U34353; NID:gl002874; PIDN:AAC44521.1; PID:gl002883
A:Experimental source: strain Fd1222
A:Note: The nucleotide sequence was submitted to the EMBL Data Library, August 1995
C:Genetics:
A:Gene: ccoH

Query Match 77.4%; Score 24; DB 2; Length 50;
Best Local Similarity 71.4%; Pred. No. 48;
Matches 5; Conservative 1; Mismatches 1; Indels 1; Gaps 0;

Qy 1 ALNVLM 7
Db 23 AVNVFM 29

RESULT 4
yaeE protein - Escherichia coli (strain K-12)
C:Species: Escherichia coli
C>Date: 12-Sep-1997 #sequence_revision 17-Sep-1997 #text_change 01-Mar-2002
C:Accession: G64767; S41304
R:Blattner, F.R.; Plunkett III, G.; Bloch, C.A.; Perna, N.T.; Burland, V.; Riley, M.; Co
.A.; Rose, D.J.; Mau, B.; Shao, Y.
Science 277, 1453-1462, 1997
A:Title: The complete genome sequence of Escherichia coli K-12.
A:Reference number: A64720; MUID:97426617; PMID:9278503
A:Accession: G64767
A:Status: nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-94 <BLAT>
A:Cross-references: GB:AE000145; GB:U00096; NID:gl786580; PIDN:AAC73494.1; PID:gl786590;
A:Experimental source: strain K-12, substrain MG1655
R:Ryder, L.; Sharples, G.J.; Lloyd, R.G.
submitted to the EMBL Data Library, December 1993
A:Description: Molecular and functional analysis of Tn1000 insertions in the aroLM-scbDQ
A:Reference number: S41303
A:Accession: S41304
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-76, 'TVSFTCKLPNPLCAAICNSSPSP' <RYD>
A:Cross-references: EMBL:X76979; NID:g440401; PIDN:CAA54286.1; PID:g440404
C:Genetics:
A:Gene: yaeE

Query Match 77.4%; Score 24; DB 2; Length 94;
Best Local Similarity 83.3%; Pred. No. 96;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ALNVLM 6
Db 51 ALNVLL 56
```

```
RESULT 5
A90684
yaeE protein [similarity] - Escherichia coli (strain O157:H7, substrain R1MD 0509952)
C:Species: Escherichia coli
C>Date: 18-Jul-2001 #sequence_revision 18-Jul-2001 #text_change 02-Nov-2001
C:Accession: A90684
R:Hayashi, T.; Makino, K.; Ohnishi, M.; Kurokawa, K.; Ishii, K.; Yokoyama, K.; Han, C.G.;
Sawawara, N.; Yasunaga, T.; Kuhara, S.; Shiba, T.; Hattori, M.; Shinagawa, H.
DNA Res. 8, 11-22, 2001
A:Title: Complete genome sequence of enterohemorrhagic Escherichia coli O157:H7 and genor
A:Reference number: A99629; MUID:21156231; PMID:11258796
A:Accession: A90684
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-94 <HAY>
A:Cross-references: GB:BA000007; PIDN:BA833864.1; PID:gl3359898; GSPDB:GN00154
A:Experimental source: strain O157:H7, substrain R1MD 0509952
C:Genetics:
A:Gene: ECs0441

Query Match 77.4%; Score 24; DB 2; Length 94;
Best Local Similarity 83.3%; Pred. No. 96;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ALNVLM 6
Db 51 ALNVLL 56

RESULT 6
E85534
yaeE protein [similarity] - Escherichia coli (strain O157:H7, substrain EDL933)
C:Species: Escherichia coli
C>Date: 16-Feb-2001 #sequence_revision 16-Feb-2001 #text_change 02-Nov-2001
C:Accession: E85534
R:Perna, N.T.; Plunkett III, G.; Burland, V.; Mau, B.; Glaesner, J.D.; Rose, D.J.; Mayhew,
iller, L.; Grotbeck, E.J.; Davis, N.W.; Lim, A.; Dimalanta, E.; Potamousis, K.; Apodaca,
Nature 409, 529-533, 2001
A:Title: Genome sequence of enterohemorrhagic Escherichia coli O157:H7.
A:Reference number: A85480; MUID:21074935; PMID:11206551
A:Accession: E85534
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-94 <STO>
A:Cross-references: GB:AE005174; NID:gl2513234; PIDN:AA654737.1; GSPDB:GN00145; UWGP:204E
A:Experimental source: strain O157:H7, substrain EDL933
C:Genetics:
A:Gene: yaeE

Query Match 77.4%; Score 24; DB 2; Length 94;
Best Local Similarity 83.3%; Pred. No. 96;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ALNVLM 6
Db 51 ALNVLL 56

RESULT 7
S24920
NADH2 dehydrogenase (ubiquinone) (EC 1.6.5.3) chain 4L - finback whale mitochondrion
C:Species: mitochondrion Balaenoptera physalus (finback whale, common rorqual)
C>Date: 31-Dec-1993 #sequence_revision 31-Dec-1993 #text_change 03-Jun-2002
C:Accession: A58851; S24920
R:Aranson, U.; Gullberg, A.; Widegren, B.
J. Mol. Evol. 33, 556-568, 1991
A:Title: The complete nucleotide sequence of the mitochondrial DNA of the fin whale, Bal
A:Reference number: A58850; MUID:92139449; PMID:1779436
A:Accession: A58851
A:Molecule type: DNA
A:Residues: 1-98 <ARN>
```

A;Cross-references: GB:X61145; NID:g12772; PIDN:CAA43447.1; PID:g12779  
 A;Note: submitted to the EMBL Data Library, August 1991

C;Genetics:

A;Genome: mitochondrion

A;Genetic code: SGC1

C;Superfamily: NADH dehydrogenase (ubiquinone) chain 4L

C;Keywords: membrane-associated complex; mitochondrion; NAD; oxidative phosphorylation;

Query Match 77.4%; Score 24; DB 2; Length 98;  
 Best Local Similarity 66.7%; Pred. No. 1e+02;  
 Matches 4; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Qy 2 LNVLMA 7

Db 6 MNILMA 11

RESULT 8

T11449

NADH2 dehydrogenase (ubiquinone) (EC 1.6.5.3) chain 4L - nine-banded armadillo mitochondrion  
 C;Species: mitochondrion Dasypus novemcinctus (nine-banded armadillo)

C;Date: 16-Jul-1999 #sequence\_revision 16-Jul-1999 #text\_change 03-Jun-2002

C;Accession: T11449

R;Arnason, U.; Gullberg, A.; Janke, A.

Mol. Biol. Evol. 14, 762-768, 1997

A;Title: Phylogenetic analysis of mitochondrial DNA suggest a sister group relationship

A;Reference number: Z1272; MUID: 97357423; PMID: 9214749

A;Accession: T11449

A;Status: preliminary; translated from GB/EMBL/DBJ

A;Molecule type: DNA

A;Residues: 1-98 <ARN>

A;Cross-references: EMBL:Y11832; NID:g2252500; PIDN:CAA72524.1; PID:g2252509

C;Genetics:

A;Genome: mitochondrion

A;Genetic code: SGC1

A;Note: NADH4L

C;Superfamily: NADH dehydrogenase (ubiquinone) chain 4L

C;Keywords: membrane-associated complex; mitochondrion; NAD; oxidative phosphorylation;

Query Match 77.4%; Score 24; DB 2; Length 98;  
 Best Local Similarity 66.7%; Pred. No. 1e+02;  
 Matches 4; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Qy 2 LNVLMA 7

Db 6 LNIIMA 11

RESULT 9

E95849

hypothetical protein [imported] - Sinorhizobium meliloti (strain 1021) megaplasmid pSymB  
 C;Species: Sinorhizobium meliloti

C;Date: 24-Aug-2001 #sequence\_revision 24-Aug-2001 #text\_change 30-Sep-2001

C;Accession: E95849

R;Finan, T.M.; Weidner, S.; Wong, K.; Buhrmester, J.; Chain, P.; Vorholter, F.J.; Hernan

Proc. Natl. Acad. Sci. U.S.A. 98, 9889-9894, 2001

A;Title: The complete sequence of the 1.683-Kb pSymB megaplasmid from the N2-fixing endo

A;Reference number: A95842; MUID: 21396508; PMID: 11481431

A;Accession: E95849

A;Status: preliminary

A;Molecule type: DNA

A;Residues: 1-51 <KUR>

A;Cross-references: GB:AL591985; PIDN:CAC48461.1; PID:g15139933; GSPDB:GN00167

A;Experimental source: strain 1021, megaplasmid pSymB

R;Galibert, F.; Finan, T.M.; Long, S.R.; Puhler, A.; Abola, P.; Ampe, F.; Barloy-Hubler,

Pels, D.; Chain, P.; Cowie, A.; Davis, R.W.; Dreano, S.; Federspiel, N.A.; Fisher, R.F.;

L.; Hyman, R.W.; Jones, T.

Science 293, 668-672, 2001

A;Authors: Kahn, D.; Kahn, M.L.; Kalman, S.; Keating, D.H.; Kiss, E.; Komp, C.; Lelaure,

hebaull, P.; Vandenbol, M.; Vorholter, F.J.; Weidner, S.; Wells, D.H.; Wong, K.; Yeh, K.

A;Title: The composite genome of the legume symbiont Sinorhizobium meliloti.

A;Reference number: A96039; MUID: 21368234; PMID: 11474104

A;Contents: annotation

C;Genetics:

A;Gene: SMB20061

A;Genome: plasmid

Query Match 74.2%; Score 23; DB 2; Length 51;

Best Local Similarity 71.4%; Pred. No. 88;

Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 1 ALNVLMA 7

Db 33 ALNLLPA 39

RESULT 10

F71843

cytochrome c553 precursor - Helicobacter pylori (strain J99)

C;Species: Helicobacter pylori

A;Variety: strain J99

C;Date: 03-Dec-1999 #sequence\_revision 03-Dec-1999 #text\_change 03-Mar-2000

C;Accession: F71843

R;Alm, R.A.; Ling, L.S.L.; Moir, D.T.; King, B.L.; Brown, E.D.; Doig, P.C.; Smith, D.R.;

; Ives, C.; Gibson, R.; Merberg, D.; Mills, S.D.; Jiang, Q.; Taylor, D.E.; Vovis, G.P.;

Nature 397, 176-180, 1999

A;Title: Genomic sequence comparison of two unrelated isolates of the human gastric patho

A;Reference number: A71800; MUID: 99120557; PMID: 9923682

A;Accession: F71843

A;Molecule type: DNA

A;Residues: 1-96 <ARN>

A;Cross-references: GB:AE001542; GB:AE001439; NID:g4155739; PIDN:AAD06721.1; PID:g4155742;

A;Experimental source: strain J99

C;Genetics:

A;Gene: jhp1148

C;Superfamily: cytochrome c6; cytochrome c6 homology

C;Keywords: chromoprotein; electron transfer; heme; iron; metalloprotein

F;1-19/Domain: signal sequence #status predicted <SIG>

F;20-96/Product: cytochrome c553 #status predicted <MAT>

F;20-92/Domain: cytochrome c6 homology <CfC>

F;29,32/Binding site: heme (Cys) (covalent) #status predicted

F;33/Binding site: heme iron (His) (axial ligand) #status predicted

Query Match 74.2%; Score 23; DB 1; Length 96;

Best Local Similarity 100.0%; Pred. No. 1.8e+02;

Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 3 NVLMA 7

Db 15 NVLMA 19

RESULT 11

S00119

hypothetical protein C - Proteus vulgaris

C;Species: Proteus vulgaris

C;Date: 31-Dec-1990 #sequence\_revision 31-Dec-1990 #text\_change 14-Apr-2003

C;Accession: S00119

R;Cole, S.T.

Eur. J. Biochem. 167, 481-488, 1987

A;Title: Nucleotide sequence and comparative analysis of the frd operon encoding the fun

d-linked ampC cephalosporinase gene.

A;Reference number: S00107; MUID: 88004470; PMID: 3308458

A;Accession: S00119

A;Molecule type: DNA

A;Residues: 1-96 <COL>

A;Cross-references: EMBL:X06151; NID:g45907; PIDN:CAA29511.1; PID:g45910

C;Superfamily: [NiFe]-hydrogenase maturation chaperone

Query Match 74.2%; Score 23; DB 2; Length 96;

Best Local Similarity 83.3%; Pred. No. 1.8e+02;

Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 2 LNVLMA 7

Db 69 LNALMA 74

Tue Feb 17 11:55:56 2004

us-09-900-147-7.rpr

```

A:Gene: AKI19
A:Introns: 30/1
C:Superfamily: unassigned Ser/Thr or Tyr-specific protein kinases; protein kinase homolog
C:Keywords: ATP; phosphotransferase; protein kinase
F:1-56/Domain: protein kinase homology (fragment) <KIN>

Query Match      71.0%; Score 22; DB 2; Length 56;
Best Local Similarity 80.0%; Pred. No. 1.7e+02;
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY      3 NVLMA 7
      |||||
Db      4 NILMA 8

RESULT 15
C34514
MFC class II histocompatibility antigen, 2 - baboon (fragment)
C:Species: Papio sp. (baboon)
C:Date: 22-Jun-1990 #sequence_revision 22-Jun-1990 #text_change 12-Jul-1996
C:Accession: C34514
R:Gyllenstein, U.B.; Erlich, H.A.
Proc. Natl. Acad. Sci. U.S.A. 86, 9986-9990, 1989
A:Title: Ancient roots for polymorphism at the HLA-DQalpha locus in primates.
A:Reference number: A34512; MUID:9009387; PMID:2513578
A:Accession: C34514
A:Status: preliminary; nucleic acid sequence not shown; not compared with conceptual tra
A:Molecule type: DNA
A:Residues: 1-63 <GYL>
C:Superfamily: class II histocompatibility antigen; immunoglobulin homology

Query Match      71.0%; Score 22; DB 2; Length 63;
Best Local Similarity 80.0%; Pred. No. 2e+02; 0; Indels 0; Gaps 0;
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY      2 LNVLM 6
      |||||
Db      56 LNVLM 60

RESULT 16
AI0225
Probable membrane protein YP01852 [imported] - Yersinia pestis (strain CO92)
C:Species: Yersinia pestis
C:Date: 02-Nov-2001 #sequence_revision 02-Nov-2001 #text_change 02-Nov-2001
C:Accession: AI0225
R:Parkhill, J.; Wren, B.W.; Thomson, N.R.; Tibball, R.W.; Holden, M.T.G.; Prentice, M.B.
deno-Tarraga, A.M.; Chillingworth, T.; Cronin, A.; Davies, R.M.; Davis, P.; Dougan, G.;
il, M.; Rutherford, K.; Simmonds, M.; Skelton, J.; Stevens, K.; Whitehead, S.; Barrell,
Nature 413, 523-527, 2001
A:Title: Genome sequence of Yersinia pestis, the causative agent of plague.
A:Reference number: AB0001; MUID:21470413; PMID:11586360
A:Accession: AI0225
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-65 <KUR>
A:Cross-references: GB:AL590842; PIDN:CAC90669.1; PID:gl5979874; GSPDB:GN00175
C:Genetics:
A:Gene: YP01852

Query Match      71.0%; Score 22; DB 2; Length 65;
Best Local Similarity 100.0%; Pred. No. 2e+02; 0; Indels 0; Gaps 0;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 ALNVL 5
      |||||
Db      49 ALNVL 53

RESULT 17
AI2345
Hypothetical protein asl4328 [imported] - Nostoc sp. (strain PCC 7120)
C:Species: Nostoc sp. PCC 7120

hemolysin - Proteus vulgaris (fragment)
S03819
C:Species: Proteus vulgaris
C:Date: 19-Mar-1997 #sequence_revision 29-Aug-1997 #text_change 16-Jul-1999
C:Accession: S03819
R:Koronakis, V.; Hughes, C.
Mol. Gen. Genet. 213, 99-104, 1988
A:Title: Identification of the promoters directing in vivo expression of hemolysin genes
A:Reference number: S03819; MUID:89127151; PMID:3065612
A:Accession: S03819
A:Status: translation not shown
A:Molecule type: DNA
A:Residues: 1-54 <KOR>
A:Cross-references: EMBL:X12571; NID:g45905; PIDN:CAA31083.1; PID:g45906
C:Genetics:
A:Gene: hlyC
C:Superfamily: hemolysin C
C:Keywords: exotoxin; hemolysis

Query Match      71.0%; Score 22; DB 2; Length 54;
Best Local Similarity 71.4%; Pred. No. 1.7e+02; 1; Indels 0; Gaps 0;
Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY      1 ALNVLM 7
      |||||
Db      33 AINVLP 39

RESULT 13
A97862
proline/betaine transporter RCL297 homolog [imported] - Rickettsia conorii (strain Malles
C:Species: Rickettsia conorii
C:Date: 30-Sep-2001 #sequence_revision 30-Sep-2001 #text_change 30-Sep-2001
C:Accession: A97862
R:Ogata, H.; Audic, S.; Renesto-Audiffren, P.; Fournier, P.E.; Barbe, V.; Samson, D.; R
Science 293, 2093-2098, 2001
A:Title: Mechanisms of Evolution in Rickettsia conorii and Rickettsia prowazekii.
A:Reference number: A97700; MUID:21442074; PMID:11557893
A:Accession: A97862
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-55 <KUR>
A:Cross-references: GB:AE006914; PIDN:AAL03835.1; PID:gl5620436; GSPDB:GN00173
C:Genetics:
A:Gene: RCL297

Query Match      71.0%; Score 22; DB 2; Length 55;
Best Local Similarity 71.4%; Pred. No. 1.7e+02; 1; Indels 0; Gaps 0;
Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY      1 ALNVLM 7
      |||||
Db      4 ALNVMP 10

RESULT 14
S66332
protein kinase AKI9 (EC 2.7.1.-) - Arabidopsis thaliana (fragment)
C:Species: Arabidopsis thaliana (mouse-ear cress)
C:Date: 19-Mar-1997 #sequence_revision 25-Apr-1997 #text_change 24-Sep-1999
C:Accession: S66332; S58268
R:Thummler, F.; Kirchner, M.; Teuber, R.; Dittrich, P.
Plant Mol. Biol. 29, 551-565, 1995
A:Title: Differential accumulation of the transcripts of 22 novel protein kinase genes
A:Reference number: S66314; MUID:96123233; PMID:8534852
A:Accession: S66332
A:Molecule type: DNA
A:Residues: 1-56 <THU>
A:Cross-references: EMBL:X86964; NID:g928903; PIDN:CAA60527.1; PID:g928904
C:Genetics:
```



A;Note: Nostoc sp. strain PCC 7120 is a synonym of Anabaena sp. strain PCC 7120  
 C;Date: 14-Dec-2001 #sequence\_revision 14-Dec-2001 #text\_change 09-Dec-2002  
 C;Accession: A12346  
 R;Kaneko, T.; Nakamura, Y.; Wolk, C.P.; Kuritz, T.; Sasamoto, S.; Watanabe, A.; Iriguchi, N.; Shimpo, S.; Sugimoto, M.; Takazawa, M.; Yamada, M.; Yasuda, M.; Tabata, S.  
 DNA Res. 8, 205-213, 2001  
 A;Title: Complete Genomic Sequence of the Filamentous Nitrogen-fixing Cyanobacterium Anabaena PCC 7120  
 A;Reference number: AB1807; MUID:21595285; PMID:11759840  
 A;Accession: A12346  
 A;Status: preliminary  
 A;Molecule type: DNA  
 A;Residues: 1-68 <KUR>  
 A;Cross-references: GB:BA000019; PIDN:BA076027.1; PID:g17133464; GSPDB:GN00179  
 A;Experimental source: strain PCC 7120  
 C;Genetics:  
 A;Gene: asl4328

Query Match 71.0%; Score 22; DB 2; Length 68;  
 Best Local Similarity 71.4%; Pred. No. 2.1e+02;  
 Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 1 ALNVLM 7  
 |||:|  
 Db 45 ALELLMA 51

RESULT 18  
 A61623  
 juvenile-hormone esterase (EC 3.1.1.59) - gypsy moth (fragment)  
 C;Species: Lymantria dispar (gypsy moth)  
 C;Date: 21-Jul-1995 #sequence\_revision 28-Jul-1995 #text\_change 21-Aug-1998  
 C;Accession: A61623  
 R;Valaitis, A.P.  
 Insect Biochem. Mol. Biol. 22, 639-648, 1992  
 A;Title: Use of concanavalin A in the purification of juvenile hormone esterase from the gypsy moth  
 A;Reference number: A61623  
 A;Accession: A61623  
 A;Status: preliminary  
 A;Molecule type: protein  
 A;Residues: 1-69 <VAL>  
 C;Superfamily: alkaline phosphatase  
 C;Keywords: carboxylic ester hydrolase; glycoprotein

Query Match 71.0%; Score 22; DB 2; Length 69;  
 Best Local Similarity 66.7%; Pred. No. 2.2e+02;  
 Matches 4; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ALNVLM 6  
 |||:|  
 Db 49 AMNVVM 54

RESULT 19  
 A36817  
 ORF X protein - human papillomavirus type 41  
 C;Species: human papillomavirus type 41  
 A;Note: host Homo sapiens (man)  
 C;Date: 23-Apr-1993 #sequence\_revision 23-Apr-1993 #text\_change 08-Oct-1999  
 C;Accession: A36817  
 R;Hirt, L.; Hirsch-Behn, A.; De Villiers, E.M.  
 Virus Res. 18, 179-190, 1990  
 A;Title: Nucleotide sequence of human papillomavirus (HPV) type 41: an unusual HPV type  
 A;Reference number: A43550  
 A;Accession: A36817  
 A;Status: translation not shown  
 A;Molecule type: DNA  
 A;Residues: 1-77 <HIR>  
 A;Cross-references: EMBL:X56147; NID:g60942; PIDN:CAA39620.1; PID:g60951

Query Match 71.0%; Score 22; DB 2; Length 77;  
 Best Local Similarity 83.3%; Pred. No. 2.5e+02;  
 Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 2 LNVLMA 7  
 |||:|  
 Db 28 LNVAMA 33

RESULT 20  
 TI7912  
 hypothetical protein a409R - Chlorella virus PBCV-1  
 C;Species: Chlorella virus PBCV-1  
 C;Date: 15-Oct-1999 #sequence\_revision 15-Oct-1999 #text\_change 15-Oct-1999  
 C;Accession: TI7912  
 R;Graves, M.V.; Van Etten, J.L.  
 submitted to the EMBL Data Library, May 1999  
 A;Reference number: Z18806  
 A;Accession: TI7912  
 A;Status: preliminary; translated from GB/EMBL/DBJ  
 A;Molecule type: DNA  
 A;Residues: 1-82 <GRA>  
 A;Cross-references: EMBL:U42580; NID:g4028896; PIDN:AAC96777.1  
 A;Experimental source: specific host Chlorella strain NC64  
 C;Genetics:  
 A;Note: a409R

Query Match 71.0%; Score 22; DB 2; Length 82;  
 Best Local Similarity 66.7%; Pred. No. 2.6e+02;  
 Matches 4; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ALNVLM 6  
 |||:|  
 Db 38 SINVLM 43

## RESULT 21

D97870

oxidoreductase homolog RCI364 [imported] - Rickettsia conorii (strain Malish 7)  
 C;Species: Rickettsia conorii  
 C;Date: 30-Sep-2001 #sequence\_revision 30-Sep-2001 #text\_change 30-Sep-2001  
 C;Accession: D97870  
 R;Ogata, H.; Audic, S.; Renesto-Audiffren, P.; Fournier, P.E.; Samson, D.; Rojo, R.; Ogata, H.; Audic, S.; Renesto-Audiffren, P.; Fournier, P.E.; Samson, D.; Rojo, R.  
 Science 293, 2093-2098, 2001  
 A;Title: Mechanisms of Evolution in Rickettsia conorii and Rickettsia prowazekii.  
 A;Reference number: A97700; MUID:21442074; PMID:11557893  
 A;Accession: D97870  
 A;Status: preliminary  
 A;Molecule type: DNA  
 A;Residues: 1-86 <KUR>  
 A;Cross-references: GB:AE006914; PIDN:AAL03902.1; PID:g15620509; GSPDB:GN00173  
 C;Genetics:  
 A;Gene: RCI364

Query Match 71.0%; Score 22; DB 2; Length 86;  
 Best Local Similarity 66.7%; Pred. No. 2.8e+02;  
 Matches 4; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ALNVLM 6  
 |||:|  
 Db 38 ALNVIL 43

## RESULT 22

S39696

WyeE protein - Bacillus subtilis  
 N;Alternate names: hypothetical protein ipa-41r  
 C;Species: Bacillus subtilis  
 C;Date: 07-Oct-1994 #sequence\_revision 26-May-1995 #text\_change 15-Oct-1999  
 C;Accession: S39696; H70052  
 R;Glaser, P.; Kunst, F.; Arnaud, M.; Coudart, M.P.; Gonzales, W.; Hullo, M.F.; Ionescu, I.; Rapoport, G.; Danchin, A.  
 Mol. Microbiol. 10, 371-384, 1993  
 A;Title: Bacillus subtilis genome project: cloning and sequencing of the 97 kb region for WyeE protein  
 A;Reference number: S39655; MUID:95020537; PMID:7934828  
 A;Accession: S39696  
 A;Status: preliminary; nucleic acid sequence not shown; translation not shown

C;Date: 30-Sep-2001 #sequence\_revision 30-Sep-2001 #text\_change 30-Sep-2001

C;Accession: D97717  
 R;Ogata, H.; Audic, S.; Renesto-Audiffren, P.; Fournier, P.E.; Barbe, V.; Samson, D.; Ro Science 293, 2093-2098, 2001  
 A;Title: Mechanisms of Evolution in Rickettsia conorii and Rickettsia prowazekii.  
 A;Reference number: A97700; MUID:21442074; PMID:11557893  
 A;Accession: D97717  
 A;Status: preliminary  
 A;Molecule type: DNA  
 A;Residues: 1-95 <KUR>  
 A;Cross-references: GB:AE006914; PIDN:AL02678.1; PID:gl5619183; GSPDB:GN00173  
 C;Genetics:  
 A;Gene: RC0140

Query Match 71.0%; Score 22; DB 2; Length 95;  
 Best Local Similarity 66.7%; Pred. No. 3.1e+02;  
 Matches 4; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 ALNVLM 6  
 :|::|  
 Db 29 ALNMIM 34

# RESULT 25

QXBO4L  
 NADH2 dehydrogenase (ubiquinone) (EC 1.6.5.3) chain 4L - bovine mitochondrion  
 N;Alternate names: NADH-ubiquinone oxidoreductase chain 4L  
 C;Species: mitochondrion Bos primigenius taurus (cattle)  
 C;Date: 18-Aug-1982 #sequence\_revision 18-Aug-1982 #text\_change 03-Jun-2002  
 C;Accession: A00429  
 C;Accession: A00429  
 A;Accession: A00429  
 A;Molecule type: DNA  
 A;Residues: 1-98 <AND>  
 A;Cross-references: GB:J01394; NID:g336430; PIDN:AB59276.1; PID:g336439; EMBL:V00654; N C;Genetics:  
 A;Genome: mitochondrion  
 A;Genetic code: SGCI  
 C;Superfamily: NADH dehydrogenase (ubiquinone) chain 4L  
 C;Keywords: membrane-associated complex; mitochondrion; NAD; oxidative phosphorylation;

Query Match 71.0%; Score 22; DB 1; Length 98;  
 Best Local Similarity 50.0%; Pred. No. 3.2e+02;  
 Matches 3; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 2 LNVLM 7  
 :|::|  
 Db 6 MNMMA 11

# RESULT 26

TL0980  
 NADH2 dehydrogenase (ubiquinone) (EC 1.6.5.3) chain 4L - pig mitochondrion  
 C;Species: mitochondrion Sus scrofa domestica (domestic pig)  
 C;Date: 16-Jul-1999 #sequence\_revision 16-Jul-1999 #text\_change 03-Jun-2002  
 C;Accession: T10980; T11878  
 R;Lin, C.S.; Liu, C.Y.; Sun, Y.L.; Chang, L.C.; Cheng, I.C.; Yang, P.C.; Mao, S.J.T.; Hu submitted to the EMBL Data Library, November 1997  
 A;Description: Complete nucleotide sequence of the porcine mitochondrial genome.  
 A;Reference number: Z17237  
 A;Accession: T10980  
 A;Status: preliminary; translated from GB/EMBL/DBJ  
 A;Molecule type: DNA  
 A;Residues: 1-98 <LIN>  
 A;Cross-references: EMBL:AF034253; NID:g4958951; PID:g4958960; PIDN:AAD34193.1  
 R;Uraing, B.M.

submitted to the EMBL Data Library, February 1999  
 A;Description: The complete mitochondrial DNA sequence of the pig (Sus scrofa).  
 A;Reference number: Z17370  
 A;Accession: T11878

A;Molecule type: DNA  
 A;Residues: 1-87 <GLA>  
 A;Cross-references: EMBL:X73124; NID:g413923; PIDN:CAAS1597.1; PID:g413965  
 A;Note: The nucleotide sequence was submitted to the EMBL Data Library, June 1993  
 R;Kunst, F.; Ogawara, N.; Moser, I.; Albertini, A.M.; Alloni, G.; Azevedo, V.; Berter  
 C.; Bron, S.; Brouillet, S.; Bruchci, C.V.; Caldwell, B.; Capuano, V.; Carter, N.M.; Cho  
 A.; Ehrlich, S.D.; Emmerson, P.T.; Entian, K.D.; Errington, J.; Fabret, C.; Ferrari, E.  
 Nature 390, 249-256, 1997  
 A;Authors: Foulger, D.; Fritz, C.; Fujita, M.; Fujita, Y.; Fuma, S.; Galizzi, A.; Galler  
 iech, J.; Harwood, C.R.; Henaut, A.; Hilbert, H.; Holsappel, S.; Hosono, S.; Hullo, M.F.  
 Koetter, P.; Konigstein, G.; Krogh, S.; Kumano, M.; Kurita, K.; Lapidus, A.; Lardinois,  
 A;Authors: Lauber, J.; Lazarevic, V.; Lee, S.M.; Levine, A.; Liu, H.; Maeda, S.; Maueel  
 Y, M.; Ogawa, K.; Ogiwara, C.; Oudega, B.; Park, S.H.; Parro, V.; Pohl, T.M.; Portetelle  
 Rieger, M.; Rivolta, C.; Rocha, E.; Roche, B.; Rose, M.; Sadaie, Y.; Sato, T.; Scanlon,  
 A;Authors: Schleich, S.; Schroter, R.; Scoffone, F.; Sekiguchi, J.; Sekowska, A.; Seron  
 akuchin, M.; Tamakoshi, A.; Tanaka, T.; Terpstra, P.; Tognoni, A.; Tosato, V.; Uchiyama,  
 T.; Winters, P.; Wipat, A.; Yamamoto, H.; Yamane, K.; Yasumoto, K.; Yata, K.; Yoshida, K  
 A;Authors: Yoshikawa, H.F.; Zumstein, E.; Yoshikawa, H.; Danchin, A.  
 A;Title: The complete genome sequence of the Gram-positive bacterium Bacillus subtilis.  
 A;Reference number: A69580; MUID:98044033; PMID:9384377  
 A;Accession: H70052  
 A;Status: preliminary; nucleic acid sequence not shown; translation not shown  
 A;Molecule type: DNA  
 A;Residues: 1-87 <KUN>  
 A;Cross-references: GB:Z99123; GB:AL009126; NID:g2636240; PIDN:CAB15839.1; PID:el186312;  
 A;Experimental source: strain 168  
 C;Genetics:  
 A;Gene: ywce  
 C;Keywords: transmembrane protein

Query Match 71.0%; Score 22; DB 2; Length 87;  
 Best Local Similarity 57.1%; Pred. No. 2.8e+02;  
 Matches 4; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 1 ALNVLM 7  
 :|::|  
 Db 62 AVNVIVA 58

# RESULT 23

T14135  
 NADH2 dehydrogenase (ubiquinone) (EC 1.6.5.3) chain 3 - white-footed mouse mitochondrion  
 C;Species: mitochondrion Peromyscus leucopus (white-footed mouse)  
 C;Date: 20-Sep-1999 #sequence\_revision 20-Sep-1999 #text\_change 03-Jun-2002  
 C;Accession: T14135  
 R;Hogan, K.M.; Davis, S.K.; Greenbaum, I.F.  
 submitted to the EMBL Data Library, November 1995  
 A;Description: Mitochondrial DNA analysis of the systematic relationships within the Per  
 A;Reference number: Z17885  
 A;Accession: T14135  
 A;Status: preliminary; translated from GB/EMBL/DBJ  
 A;Molecule type: DNA  
 A;Residues: 1-89 <HOG>  
 A;Cross-references: EMBL:U40252; NID:gl655748; PID:gl655749; PIDN:AAB17918.1  
 C;Genetics:  
 A;Genome: mitochondrion

A;Note: ND3  
 C;Superfamily: NADH dehydrogenase (ubiquinone) chain 3  
 C;Keywords: membrane-associated complex; mitochondrion; NAD; oxidative phosphorylation;

Query Match 71.0%; Score 22; DB 2; Length 89;  
 Best Local Similarity 66.7%; Pred. No. 2.9e+02;  
 Matches 4; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 2 LNVLM 7  
 :|::|  
 Db 1 MNMLMA 6

# RESULT 24

D97717  
 hypothetical protein RC0140 [imported] - Rickettsia conorii (strain Malish 7)  
 C;Species: Rickettsia conorii

A;Status: preliminary; translated from GB/EMBL/DBJ  
 A;Molecule type: DNA  
 A;Residues: 1-98 <URS>  
 A;Cross-references: EMBL:AJ002189; PIDN:CAA05240.1  
 C;Genetics:

A;Gene: NADH4L

A;Genome: mitochondrion

A;Genetic code: SGC1

C;Superfamily: NADH dehydrogenase (ubiquinone) chain 4L

C;Keywords: membrane-associated complex; mitochondrion; NAD; oxidative phosphorylation;

Query Match 71.0%; Score 22; DB 2; Length 98;  
 Best Local Similarity 50.0%; Pred. No. 3.2e+02;  
 Matches 3; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

Qy 2 LNVLMA 7

Db 6 MNILMA 11

# RESULT 27

T11058

NADH2 dehydrogenase (ubiquinone) (EC 1.6.5.3) chain 4L - sheep mitochondrion

C;Species: mitochondrion Ovis orientalis aries, Ovis ammon aries (domestic sheep)

C;Date: 16-Jul-1999 #sequence\_revision 16-Jul-1999 #text\_change 03-Jun-2002

C;Accession: T11058

R;Händler, S.; Lewalaki, H.; Wassmuth, R.; Janke, A.

J. Mol. Evol. 47, 441-448, 1998

A;Title: The complete mitochondrial DNA sequence of the domestic sheep (Ovis aries) and

A;Reference number: Z17245; MUID:98440761; PMID:9767689

A;Accession: T11058

A;Status: preliminary; translated from GB/EMBL/DBJ

A;Molecule type: DNA

A;Residues: 1-98 <HIE>

A;Cross-references: EMBL:AF010406; NID:G3445513; PID:G3366630; PIDN:AAD10103.1

A;Experimental source: strain Merinolandschaf; liver

C;Genetics:

A;Genome: mitochondrion

A;Genetic code: SGC1

A;Note: ND4L

C;Superfamily: NADH dehydrogenase (ubiquinone) chain 4L

C;Keywords: membrane-associated complex; mitochondrion; NAD; oxidative phosphorylation;

Query Match 71.0%; Score 22; DB 2; Length 98;  
 Best Local Similarity 50.0%; Pred. No. 3.2e+02;  
 Matches 3; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

Qy 2 LNVLMA 7

Db 6 MNILMA 11

# RESULT 28

JC5794

elongin C - fruit fly (Drosophila melanogaster)

C;Species: Drosophila melanogaster

C;Date: 24-Jan-1998 #sequence\_revision 13-Mar-1998 #text\_change 17-Nov-2000

C;Accession: JC5794

R;Abo, T.; Conrad, M.N.

Biochem. Biophys. Res. Commun. 241, 334-340, 1997

A;Title: Molecular cloning of DNAs encoding the regulatory subunits of elongin from Sac

A;Reference number: JC5792; MUID:98086352; PMID:9425272

A;Accession: JC5794

A;Status: nucleic acid sequence not shown

A;Molecule type: mRNA

A;Residues: 1-98 <ASO>

A;Cross-references: DDBJ:AB007693; NID:G2780364; PIDN:BA24287.1; PID:G2780365

C;Comment: This protein binds to each other in the absence of elongin A to form a binary

or induction of elongin A activity.

C;Genetics:

A;Cross-references: FlyBase:FBgn0023211

C;Superfamily: elongin C

Query Match 71.0%; Score 22; DB 2; Length 98;  
 Best Local Similarity 71.4%; Pred. No. 3.2e+02;  
 Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 1 ALNVLMA 7

Db 86 ALNLMA 92

# RESULT 29

C70211

hypothetical protein BBA35 - Lyme disease spirochete plasmid A/lp54

C;Species: Borrelia burgdorferi (Lyme disease spirochete)

C;Date: 13-Feb-1998 #sequence\_revision 13-Feb-1998 #text\_change 08-Oct-1999

C;Accession: C70211

R;Fraser, C.M.; Casjens, S.; Huang, W.M.; Sutton, G.G.; Clayton, R.; Lathigra, R.; White

son, D.; Peterson, J.; Kervavage, A.R.; Quackenbush, J.; Salzberg, S.; Hanson, M.; Vugt,

; Bowman, C.; Garland, S.; Fujii, C.; Cotton, M.D.; Horst, K.; Roberts, K.; Hatch, B.

Nature 390, 580-586, 1997

A;Authors: Smith, H.O.; Venter, J.C.

A;Title: Genomic sequence of a Lyme disease spirochaete, Borrelia burgdorferi.

A;Reference number: A70100; MUID:98065943; PMID:9403685

A;Accession: C70211

A;Status: preliminary; nucleic acid sequence not shown; translation not shown

A;Molecule type: DNA

A;Residues: 1-36 <KLE>

A;Cross-references: GB:AE000790; NID:G2690224; PIDN:AAC66298.1; PID:G2690297; TIGR:BBA35

A;Experimental source: strain B31

C;Genetics:

A;Genome: plasmid

Query Match 67.7%; Score 21; DB 2; Length 36;  
 Best Local Similarity 66.7%; Pred. No. 1.9e+02;  
 Matches 4; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ALNVLM 6

Db 8 ALNLLL 13

# RESULT 30

PN0624

alpha-interneixin - pig (fragments)

C;Species: Sus scrofa domestica (domestic pig)

C;Date: 05-Aug-1994 #sequence\_revision 05-Aug-1994 #text\_change 17-Mar-1999

C;Accession: PN0624

R;Tanaka, J.; Ogawara, M.; Ando, S.; Shibata, M.; Yatani, R.; Kusagawa, M.; Inagaki, M.

Biochem. Biophys. Res. Commun. 196, 115-123, 1993

A;Title: Phosphorylation of a 62kD porcine alpha-interneixin, a newly identified intermed

A;Reference number: PN0624; MUID:94029981; PMID:8216281

A;Accession: PN0624

A;Molecule type: protein

A;Residues: 1-18,19-28;29-45;46-57 <TAN>

C;Comment: This intermediate filament protein polymerizes to form filament itself, in vi

C;Superfamily: cytoskeletal keratin

C;Keywords: brain; intermediate filament

Query Match 67.7%; Score 21; DB 2; Length 57;  
 Best Local Similarity 83.3%; Pred. No. 3.1e+02;  
 Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 2 LNVLMA 7

Db 42 LNVKMA 47

# RESULT 31

T16930

hypothetical protein T23F2.5 - Caenorhabditis elegans

C;Species: Caenorhabditis elegans

C;Date: 20-Sep-1999 #sequence\_revision 20-Sep-1999 #text\_change 02-Jun-2000

C;Accession: T16930

R;Du, Z.

submitted to the EMBL Data Library, October 1995  
A;Description: The sequence of C. elegans cosmid T23F2.  
A;Reference number: Z18608  
A;Accession: T16930  
A;Status: preliminary; translated from GB/EMBL/DBJ  
A;Molecule type: DNA  
A;Residues: 1-57 <DUZ>  
A;Cross-references: EMBL:U39649; NID:g1049370; PID:g1049375; PIDN:AAA80387.1; CESP:T23F2  
C;Genetics:  
A;Gene: CESP:T23F2.5  
A;Introns: 18/3  
C;Superfamily: conserved hypothetical protein b2666

Query Match 67.7%; Score 21; DB 2; Length 57;  
Best Local Similarity 50.0%; Pred. No. 3.1e+02;  
Matches 3; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 1 ALNVL 6  
|:|:  
33 AINILL 38

RESULT 32  
A69337  
conserved hypothetical protein AF0697 - Archaeoglobus fulgidus  
C;Species: Archaeoglobus fulgidus  
C;Date: 05-Dec-1997 #sequence\_revision 05-Dec-1997 #text\_change 29-Sep-1999  
C;Accession: A69337  
R;Klenk, H.P.; Clayton, R.A.; Tomb, J.P.; White, O.; Nelson, K.E.; Ketchum, K.A.; Dodson  
J.; Fleischmann, R.D.; Quackenbush, J.; Lee, N.H.; Sutton, G.G.; Gill, S.; Kirkness, E.F.  
Glodek, A.; Zhou, L.; Overbeek, R.; Gocayne, J.D.; Weidman, J.F.; McDonald, L.  
Nature 390, 364-370, 1997  
A;Authors: Utterback, T.; Cotton, M.D.; Spriggs, T.; Artach, P.; Kaine, B.P.; Sykes, S.  
Smith, H.O.; Woese, C.R.; Ventet, J.C.  
A;Title: The complete genome sequence of the hyperthermophilic, sulfate-reducing archae  
A;Reference number: A69250; MUID:98049343; PMID:9389475  
A;Accession: A69337  
A;Status: preliminary; nucleic acid sequence not shown; translation not shown  
A;Molecule type: DNA  
A;Residues: 1-61 <MLE>  
A;Cross-references: GB:AE001056; GB:AE000782; NID:g2689379; PIDN:AAB90542.1; PID:g264991  
C;Superfamily: hypothetical protein MJ0546

Query Match 67.7%; Score 21; DB 2; Length 61;  
Best Local Similarity 66.7%; Pred. No. 3.4e+02;  
Matches 4; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 ALNVL 6  
|:|:  
17 ALNLL 22

RESULT 33  
H90118  
probable small nuclear ribonucleoprotein G [imported] - Guillardia theta nucleomorph  
C;Species: nucleomorph Guillardia theta  
A;Note: a nucleomorph is the vestigial nucleus of a eukaryotic endosymbiont  
C;Date: 10-May-2001 #sequence\_revision 10-May-2001 #text\_change 24-May-2001  
C;Accession: H90118  
R;Douglas, S.; Zauner, S.; Fraunholz, M.; Beaton, M.; Penny, S.; Deng, L.T.; Wu, X.; Rei  
Nature 410, 1091-1096, 2001  
A;Title: The highly reduced genome of an enslaved algal nucleus.  
A;Reference number: A99082; MUID:11323671; PMID:11323671  
A;Accession: H90118  
A;Status: preliminary  
A;Molecule type: DNA  
A;Residues: 1-63 <DOU>  
A;Cross-references: GB:AF083031; NID:g13794269; PIDN:AAK39646.1; GSPDB:GN00152  
C;Genetics:  
A;Gene: snrpg  
A;Map position: 3  
A;Genome: nucleomorph  
C;Keywords: nucleomorph

Query Match 67.7%; Score 21; DB 2; Length 63;  
Best Local Similarity 80.0%; Pred. No. 3.5e+02;  
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 2 LNVL 6  
|:|:  
23 MNVLM 27

RESULT 34  
I38244  
gene SPHAR protein - human  
C;Species: Homo sapiens (man)  
C;Date: 29-May-1998 #sequence\_revision 29-May-1998 #text\_change 05-Nov-1999  
C;Accession: I38244  
R;Digweed, M.; Gunther, U.; Schneider, R.; Seyschab, H.; Friedl, R.; Sperling, K.  
Mol. Cell. Biol. 15, 305-314, 1995  
A;Title: Irreversible repression of DNA synthesis in Fanconi anemia cells is alleviated b  
A;Reference number: I38244; MUID:95098005; PMID:7799938  
A;Accession: I38244  
A;Status: preliminary; translated from GB/EMBL/DBJ  
A;Molecule type: DNA  
A;Residues: 1-63 <RES>  
A;Cross-references: EMBL:X82554; NID:g575271; PIDN:CAAS7901.1; PID:g575272  
C;Genetics:  
A;Gene: SPHAR

Query Match 67.7%; Score 21; DB 2; Length 63;  
Best Local Similarity 80.0%; Pred. No. 3.5e+02;  
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 ALNVL 5  
|:|:  
22 ALNIL 26

RESULT 35  
H90804  
hypothetical protein ECs1408 [imported] - Escherichia coli (strain O157:H7, substrain R1  
C;Species: Escherichia coli  
C;Date: 18-Jul-2001 #sequence\_revision 18-Jul-2001 #text\_change 18-Jul-2001  
C;Accession: H90804  
R;Hayashi, T.; Makino, K.; Ohnishi, M.; Kurokawa, K.; Ishii, K.; Yokoyama, K.; Han, C.G.  
gasawara, N.; Yasunaga, T.; Kuhara, S.; Shiba, T.; Hattori, M.; Shinagawa, H.  
DNA Res. 8, 11-22, 2001  
A;Title: Complete genome sequence of enterohemorrhagic Escherichia coli O157:H7 and geno  
A;Reference number: A99629; MUID:21156231; PMID:11258796  
A;Accession: H90804  
A;Status: preliminary  
A;Molecule type: DNA  
A;Residues: 1-65 <HAY>  
A;Cross-references: GB:BA000007; PIDN:BA034831.1; PID:gl3360868; GSPDB:GN00154  
A;Experimental source: strain O157:H7, substrain RMD 050952  
C;Genetics:  
A;Gene: ECs1408

Query Match 67.7%; Score 21; DB 2; Length 65;  
Best Local Similarity 57.1%; Pred. No. 3.6e+02;  
Matches 4; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 1 ALNVL 7  
|:|:  
9 ALDIL 15

RESULT 36  
E91196  
hypothetical protein ECs4541 [imported] - Escherichia coli (strain O157:H7, substrain R1  
C;Species: Escherichia coli  
C;Date: 18-Jul-2001 #sequence\_revision 18-Jul-2001 #text\_change 18-Jul-2001  
C;Accession: E91196  
R;Hayashi, T.; Makino, K.; Ohnishi, M.; Kurokawa, K.; Ishii, K.; Yokoyama, K.; Han, C.G

```

gasawara, N.; Yasunaga, T.; Kuhara, S.; Shiba, T.; Hattori, M.; Shinagawa, H.
DNA Res. 8, 11-22, 2001
A;Title: Complete genome sequence of enterohemorrhagic Escherichia coli O157:H7 and geno
A;Reference number: A99629; MUID:21156231; PMID:11258796
A;Accession: E91196
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-65 <HAY>
A;Cross-references: GB:BA000007; PIDN:BAB37964.1; PID:gl3364016; GSPDB:GN00154
A;Experimental source: strain O157:H7, substrain RIMD 050952
C;Genetics:
A;Gene: EC64541

Query Match      67.7%; Score 21; DB 2; Length 65;
Best Local Similarity 57.1%; Pred. No. 3.6e+02;
Matches 4; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ALNVLMA 7
Db 9 ALDILIA 15

RESULT 37
D86043
unknown protein encoded within prophage CP-933L [imported] - Escherichia coli (strain O1
C;Species: Escherichia coli
C;Date: 16-Feb-2001 #sequence_revision 16-Feb-2001 #text_change 14-Sep-2001
C;Accession: D86043
R;Perna, N.T.; Plunkett III, G.; Burland, V.; Mau, B.; Glasner, J.D.; Rose, D.J.; Mayhew
iller, L.; Grotbeck, E.J.; Davis, N.W.; Lim, A.; Dimalanta, E.; Potamouis, K.; Apodaca,
Nature 409, 529-533, 2001
A;Title: Genome sequence of enterohemorrhagic Escherichia coli O157:H7.
A;Reference number: A85480; MUID:21074935; PMID:11206551
A;Accession: D86043
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-65 <STO>
A;Cross-references: GB:AE005174; NID:gl2518429; PIDN:AAG58808.1; GSPDB:GN00145; UWGP:Z50
A;Experimental source: strain O157:H7, substrain EDL933
C;Genetics:
A;Gene: Z5093

Query Match      67.7%; Score 21; DB 2; Length 65;
Best Local Similarity 57.1%; Pred. No. 3.6e+02;
Matches 4; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ALNVLMA 7
Db 9 ALDILIA 15

RESULT 38
E85613
hypothetical protein Z1225 [imported] - Escherichia coli (strain O157:H7, substrain EDL9
C;Species: Escherichia coli
C;Date: 16-Feb-2001 #sequence_revision 16-Feb-2001 #text_change 14-Sep-2001
C;Accession: E85613
R;Perna, N.T.; Plunkett III, G.; Burland, V.; Mau, B.; Glasner, J.D.; Rose, D.J.; Mayhew
iller, L.; Grotbeck, E.J.; Davis, N.W.; Lim, A.; Dimalanta, E.; Potamouis, K.; Apodaca,
Nature 409, 529-533, 2001
A;Title: Genome sequence of enterohemorrhagic Escherichia coli O157:H7.
A;Reference number: A85480; MUID:21074935; PMID:11206551
A;Accession: E85613
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-65 <STO>
A;Cross-references: GB:AE005174; NID:gl2514041; PIDN:AAG55369.1; GSPDB:GN00145; UWGP:Z12
A;Experimental source: strain O157:H7, substrain EDL933
C;Genetics:
A;Gene: Z1225

Query Match      67.7%; Score 21; DB 2; Length 65;
Best Local Similarity 57.1%; Pred. No. 3.6e+02;

```

```

Matches 4; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ALNVLMA 7
Db 9 ALDILIA 15

RESULT 39
H98064
hypothetical protein spr1545 [imported] - Streptococcus pneumoniae (strain R6)
C;Species: Streptococcus pneumoniae
C;Date: 22-Oct-2001 #sequence_revision 22-Oct-2001 #text_change 22-Oct-2001
C;Accession: H98064
R;Hoskins, J.A.; Alborn Jr., W.; Arnold, J.; Blaszcak, L.; Burgett, S.; DeHoff, B.S.; E
e, R.; LeBlanc, D.J.; Lee, L.N.; Lefkowitz, E.J.; Lu, J.; Matsushima, P.; McAhren, S.; M
y, P.; Sun, P.M.; Winkler, M.E.
J. Bacteriol. 183, 5709-5717, 2001
A;Authors: Yang, Y.; Young-Bellido, M.; Zhao, G.; Zook, C.; Baltz, R.H.; Jaskunas, S.R.;
A;Title: Genome of the Bacterium Streptococcus pneumoniae Strain R6.
A;Reference number: A97872; MUID:21429245; PMID:11544234
A;Accession: H98064
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-66 <KUR>
A;Cross-references: GB:AE007317; PIDN:AAL00349.1; PID:gl5459210; GSPDB:GN00174
C;Genetics:
A;Gene: spr1545

Query Match      67.7%; Score 21; DB 2; Length 66;
Best Local Similarity 66.7%; Pred. No. 3.7e+02;
Matches 4; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Qy 2 LNVILMA 7
Db 26 LSVILMA 31

RESULT 40
AH0390
hypothetical protein YP03217 [imported] - Yersinia pestis (strain CO92)
C;Species: Yersinia pestis
C;Date: 02-Nov-2001 #sequence_revision 02-Nov-2001 #text_change 02-Nov-2001
C;Accession: AH0390
R;Parkhill, J.; Wren, B.W.; Thomson, N.R.; Titball, R.W.; Holden, M.T.G.; Prentice, M.B.
deno-Tarraga, A.M.; Chillingworth, T.; Cronin, A.; Davies, R.M.; Davis, P.; Dougan, G.;
il, M.; Rutherford, K.; Simmonds, M.; Skelton, J.; Stevens, K.; Whitehead, S.; Barrell,
Nature 413, 523-527, 2001
A;Title: Genome sequence of Yersinia pestis, the causative agent of plague.
A;Reference number: AB0001; MUID:21470413; PMID:11586360
A;Accession: AH0390
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-71 <KUR>
A;Cross-references: GB:AL590842; PIDN:CAC92452.1; PID:gl5981153; GSPDB:GN00175
C;Genetics:
A;Gene: YP03217

Query Match      67.7%; Score 21; DB 2; Length 71;
Best Local Similarity 80.0%; Pred. No. 4e+02;
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 3 NVILMA 7
Db 12 NVILMA 16

RESULT 41
AC2335
hypothetical protein asl4234 [imported] - Nostoc sp. (strain PCC 7120)
C;Species: Nostoc sp. PCC 7120
A;Note: Nostoc sp. strain PCC 7120 is a synonym of Anabaena sp. strain PCC 7120
C;Date: 14-Dec-2001 #sequence_revision 14-Dec-2001 #text_change 09-Dec-2002
C;Accession: AC2335

```

search completed: February 17, 2004, 10:58:19

Job time : 5.22772 secs

**THIS PAGE BLANK (USPTO)**



GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: February 17, 2004, 10:53:48 ; Search time 10.2574 Seconds  
(without alignments)  
142.889 Million cell updates/sec

Title: US-09-900-147-7

Perfect score: 31

Sequence: 1 ALNVLMA 7

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 801455 seqs, 209382283 residues

Total number of hits satisfying chosen parameters: 318354

Minimum DB seq length: 0

Maximum DB seq length: 100

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications AA:\*

1: /cgn2\_6/ptodata/1/pubpaa/US07\_PUBCOMB.pep.\*  
2: /cgn2\_6/ptodata/1/pubpaa/PCT\_NEW\_PUB.pep.\*  
3: /cgn2\_6/ptodata/1/pubpaa/US06\_NEW\_PUB.pep.\*  
4: /cgn2\_6/ptodata/1/pubpaa/US06\_PUBCOMB.pep.\*  
5: /cgn2\_6/ptodata/1/pubpaa/US07\_NEW\_PUB.pep.\*  
6: /cgn2\_6/ptodata/1/pubpaa/PCTUS\_PUBCOMB.pep.\*  
7: /cgn2\_6/ptodata/1/pubpaa/US08\_NEW\_PUB.pep.\*  
8: /cgn2\_6/ptodata/1/pubpaa/US08\_PUBCOMB.pep.\*  
9: /cgn2\_6/ptodata/1/pubpaa/US09A\_PUBCOMB.pep.\*  
10: /cgn2\_6/ptodata/1/pubpaa/US09B\_PUBCOMB.pep.\*  
11: /cgn2\_6/ptodata/1/pubpaa/US09C\_PUBCOMB.pep.\*  
12: /cgn2\_6/ptodata/1/pubpaa/US09\_NEW\_PUB.pep.\*  
13: /cgn2\_6/ptodata/1/pubpaa/US10A\_PUBCOMB.pep.\*  
14: /cgn2\_6/ptodata/1/pubpaa/US10B\_PUBCOMB.pep.\*  
15: /cgn2\_6/ptodata/1/pubpaa/US10C\_PUBCOMB.pep.\*  
16: /cgn2\_6/ptodata/1/pubpaa/US10\_NEW\_PUB.pep.\*  
17: /cgn2\_6/ptodata/1/pubpaa/US60\_NEW\_PUB.pep.\*  
18: /cgn2\_6/ptodata/1/pubpaa/US60\_PUBCOMB.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	31	100.0	7	10	US-09-900-147-7
2	31	100.0	14	10	US-09-900-147-11
3	31	100.0	16	10	US-09-900-147-5
4	31	100.0	19	10	US-09-900-147-3
5	31	100.0	19	10	US-09-900-147-16
6	31	100.0	30	10	US-09-900-147-6
7	31	100.0	37	10	US-09-900-147-1
8	31	100.0	74	15	US-10-214-188-10
9	27	87.1	19	10	US-09-900-147-15
10	27	87.1	56	14	US-10-044-359-8
11	26	83.9	41	9	US-09-820-893-58
12	26	83.9	90	9	US-09-925-299-824
13	26	83.9	90	11	US-09-925-299-824
14	25	80.6	87	9	US-09-853-161-89
15	25	80.6	87	9	US-09-852-659A-89

16	25	80.6	87	10	US-09-852-797-89	Sequence 89, Appl
17	24	77.4	27	9	US-09-864-761-35481	Sequence 35481, A
18	24	77.4	32	15	US-10-106-698-7834	Sequence 7834, Ap
19	24	77.4	64	10	US-09-738-626-4982	Sequence 4982, Ap
20	24	77.4	65	9	US-09-764-860-538	Sequence 538, App
21	24	77.4	65	12	US-10-212-872-538	Sequence 538, App
22	24	77.4	65	15	US-10-074-095-538	Sequence 538, App
23	24	77.4	71	10	US-09-925-300-1456	Sequence 1456, Ap
24	24	77.4	96	12	US-10-029-386-29049	Sequence 29049, A
25	23	74.2	9	10	US-09-900-147-2	Sequence 2, Appli
26	23	74.2	20	9	US-09-864-761-41818	Sequence 41818, A
27	23	74.2	20	10	US-09-900-147-4	Sequence 4, Appli
28	23	74.2	26	9	US-09-864-761-45731	Sequence 45731, A
29	22	71.0	9	12	US-10-024-652-27	Sequence 27, Appl
30	22	71.0	9	12	US-10-024-652-142	Sequence 142, App
31	22	71.0	9	12	US-10-024-652-348	Sequence 348, App
32	22	71.0	9	12	US-10-024-652-412	Sequence 412, App
33	22	71.0	9	12	US-10-024-652-444	Sequence 444, App
34	22	71.0	9	12	US-10-024-652-501	Sequence 501, App
35	22	71.0	9	12	US-10-024-652-626	Sequence 626, App
36	22	71.0	9	12	US-10-024-652-636	Sequence 636, App
37	22	71.0	9	12	US-10-024-652-926	Sequence 926, App
38	22	71.0	9	12	US-10-024-652-976	Sequence 976, App
39	22	71.0	9	12	US-10-024-652-977	Sequence 977, App
40	22	71.0	9	12	US-10-024-652-1076	Sequence 1076, Ap
41	22	71.0	9	12	US-10-024-652-1219	Sequence 1219, Ap
42	22	71.0	9	12	US-10-024-652-1272	Sequence 1272, Ap
43	22	71.0	9	12	US-10-024-652-1279	Sequence 1279, Ap
44	22	71.0	9	12	US-10-024-652-1311	Sequence 1311, Ap
45	22	71.0	9	12	US-10-024-652-1383	Sequence 1383, Ap

#### ALIGNMENTS

RESULT 1  
US-09-900-147-7  
; Sequence 7, Application US/09900147  
; Patent No. US20020103121A1  
; GENERAL INFORMATION:  
; APPLICANT: La Thangue, Nicholas B  
; APPLICANT: Bandara, Lisantha R  
; TITLE OF INVENTION: Peptide antagonists of DP transcription factors  
; FILE REFERENCE: 620-67  
; CURRENT FILING DATE: 2001-07-09  
; CURRENT APPLICATION NUMBER: US/09/900,147  
; PRIOR FILING DATE: EARLIER APPLICATION NUMBER: 09/308,935  
; PRIOR APPLICATION NUMBER: EARLIER FILING DATE: 1999-05-27  
; PRIOR FILING DATE: EARLIER APPLICATION NUMBER: GB 9626589.7  
; NUMBER OF SEQ ID NOS: 18  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 7  
; LENGTH: 7  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Description of Artificial Sequence: Synthetic  
US-09-900-147-7

Query Match 100.0%; Score 31; DB 10; Length 7;  
Best Local Similarity 100.0%; Pred. NO. 7e+05;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ALNVLMA 7  
| | | | |  
Db 1 ALNVLMA 7

RESULT 2  
US-09-900-147-11  
; Sequence 11, Application US/09900147  
; Patent No. US20020103121A1



```
; PRIOR FILING DATE: EARLIER FILING DATE: 1996-12-20
; NUMBER OF SEQ ID NOS: 18
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 6
; LENGTH: 30
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Synthetic
US-09-900-147-6

Query Match          100.0%; Score 31; DB 10; Length 30;
Best Local Similarity 100.0%; Pred. No. 3.5; Indels 0; Gaps 0;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ALNVLMA 7
Db 3 ALNVLMA 9

RESULT 7
US-09-900-147-1
; Sequence 1, Application US/09900147
; Patent No. US20020103121A1
; GENERAL INFORMATION:
; APPLICANT: La Thangue, Nicholas B
; TITLE OF INVENTION: Peptide antagonists of DP transcription factors
; FILE REFERENCE: 620-67
; CURRENT APPLICATION NUMBER: US/09/900,147
; CURRENT FILING DATE: 2001-07-09
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 09/308,935
; PRIOR FILING DATE: EARLIER FILING DATE: 1999-05-27
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: GB 9626589.7
; PRIOR FILING DATE: EARLIER FILING DATE: 1996-12-20
; NUMBER OF SEQ ID NOS: 18
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1
; LENGTH: 37
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Synthetic
US-09-900-147-1

Query Match          100.0%; Score 31; DB 10; Length 37;
Best Local Similarity 100.0%; Pred. No. 4.4; Indels 0; Gaps 0;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ALNVLMA 7
Db 10 ALNVLMA 16

RESULT 8
US-10-214-188-10
; Sequence 10, Application US/10214188
; Publication No. US2003002260A1
; GENERAL INFORMATION:
; APPLICANT: LA THANGUE, NICHOLAS B.
; BERNARDS, RENE
; HJUMANS, ELEANORE M.
; TITLE OF INVENTION: TRANSCRIPTION FACTOR E2F-5
; NUMBER OF SEQUENCES: 25
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: NIXON & VANDERHYE P.C.
; STREET: 1100 NORTH GLEBE ROAD
; CITY: ARLINGTON
; STATE: VIRGINIA
; COUNTRY: U.S.A.
; ZIP: 22201-4714
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/10/214,188
; FILING DATE: 08-AUG-2002
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/894,139
; FILING DATE: 13-AUG-1997
; ATTORNEY/AGENT INFORMATION:
; NAME: WILSON, MARY J.
; REGISTRATION NUMBER: 32,955
; REFERENCE/DOCKET NUMBER: 620-22
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703) 816-4000
; TELEFAX: (703) 816-4100
; INFORMATION FOR SEQ ID NO: 10:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 74 amino acids
; TYPE: amino acid
; STRANDEDNESS: <Unknown>
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; SEQUENCE DESCRIPTION: SEQ ID NO: 10:
US-10-214-188-10

Query Match          100.0%; Score 31; DB 15; Length 74;
Best Local Similarity 100.0%; Pred. No. 9.2; Indels 0; Gaps 0;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ALNVLMA 7
Db 52 ALNVLMA 58

RESULT 9
US-09-900-147-15
; Sequence 15, Application US/09900147
; Patent No. US20020103121A1
; GENERAL INFORMATION:
; APPLICANT: La Thangue, Nicholas B
; APPLICANT: Bandara, Lasantha R
; TITLE OF INVENTION: Peptide antagonists of DP transcription factors
; FILE REFERENCE: 620-67
; CURRENT APPLICATION NUMBER: US/09/900,147
; CURRENT FILING DATE: 2001-07-09
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 09/308,935
; PRIOR FILING DATE: EARLIER FILING DATE: 1999-05-27
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: GB 9626589.7
; PRIOR FILING DATE: EARLIER FILING DATE: 1996-12-20
; NUMBER OF SEQ ID NOS: 18
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 15
; LENGTH: 19
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Mutant peptide
US-09-900-147-15

Query Match          87.1%; Score 27; DB 10; Length 19;
Best Local Similarity 85.7%; Pred. No. 17;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 ALNVLMA 7
Db 7 ALNVLMA 13

RESULT 10
US-10-044-359-8
; Sequence 8, Application US/10044359
```

Tue Feb 17 11:55:56 2004

```

; Publication No. US20020160454A1
; GENERAL INFORMATION:
; APPLICANT: Herzman, Rafael
; APPLICANT: Wong, James F.
; APPLICANT: Lee, Jian-Ming
; TITLE OF INVENTION: SCORPION TOXINS
; FILE REFERENCE: BB1367 US NA
; CURRENT APPLICATION NUMBER: US/10/044,359
; CURRENT FILING DATE: 2002-01-11
; PRIOR APPLICATION NUMBER: 09/599,416
; PRIOR FILING DATE: 2000-06-22
; PRIOR APPLICATION NUMBER: 60/140,227
; PRIOR FILING DATE: 1999-06-22
; NUMBER OF SEQ ID NOS: 28
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 8
; LENGTH: 56
; TYPE: PRT
; ORGANISM: Hottentotta judaica
US-10-044-359-8

Query Match      87.1%; Score 27; DB 14; Length 56;
Best Local Similarity 71.4%; Pred. No. 54;
Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 ALNVLM 7
Db 13 AMNVMA 19

RESULT 11
US-09-820-893-58
; Sequence 58, Application US/09820893
; Patent No. US20020076705A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: 31 Human Secreted Proteins
; FILE REFERENCE: P2033P1
; CURRENT APPLICATION NUMBER: US/09/820,893
; CURRENT FILING DATE: 2001-03-30
; PRIOR APPLICATION NUMBER: 09/531,119
; PRIOR FILING DATE: 2000-03-20
; PRIOR APPLICATION NUMBER: 60/102,895
; PRIOR FILING DATE: 1998-10-02
; NUMBER OF SEQ ID NOS: 140
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 58
; LENGTH: 41
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (41)
; OTHER INFORMATION: Xaa equals stop translation
US-09-820-893-58

Query Match      83.9%; Score 26; DB 9; Length 41;
Best Local Similarity 85.7%; Pred. No. 65;
Matches 6; Conservative 0; Mismatches 1; Indels 1; Gaps 0;

QY 1 ALNVLM 7
Db 8 ALNVLM 14

RESULT 12
US-09-925-299-824
; Sequence 824, Application US/09925299
; Patent No. US20020055627A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies
; FILE REFERENCE: PA102
; PRIOR APPLICATION NUMBER: 60/265,583
; CURRENT FILING DATE: 2001-05-11
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
US-09-925-299-824

Query Match      83.9%; Score 26; DB 9; Length 90;
Best Local Similarity 83.3%; Pred. No. 1.5e+02;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 ALNVLM 6
Db 44 ALNVLM 49

RESULT 13
US-09-925-299-824
; Sequence 824, Application US/09925299
; Publication No. US20030040617A9
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies
; FILE REFERENCE: PA102
; CURRENT APPLICATION NUMBER: US/09/925,299
; CURRENT FILING DATE: 2001-08-10
; PRIOR APPLICATION NUMBER: PCT/US00/05883
; PRIOR FILING DATE: 2000-03-08
; PRIOR APPLICATION NUMBER: 60/124,270
; PRIOR FILING DATE: 1999-03-12
; NUMBER OF SEQ ID NOS: 1556
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 824
; LENGTH: 90
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (36)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
US-09-925-299-824

Query Match      83.9%; Score 26; DB 9; Length 90;
Best Local Similarity 83.3%; Pred. No. 1.5e+02;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 ALNVLM 6
Db 44 ALNVLM 49

RESULT 14
US-09-853-161-89
; Sequence 89, Application US/09853161
; Patent No. US20020076756A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: 28 Human Secreted Proteins
; FILE REFERENCE: P2003P3
; CURRENT APPLICATION NUMBER: US/09/853,161
; CURRENT FILING DATE: 2001-05-11
; PRIOR APPLICATION NUMBER: 60/265,583
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
US-09-853-161-89

```



FEATURE:  
NAME/KEY: SITE  
LOCATION: (11)  
OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids  
NAME/KEY: SITE  
LOCATION: (86)  
OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids  
US-09-852-797-89  
Query Match 80.6%; Score 25; DB 10; Length 87;  
Best Local Similarity 71.4%; Pred. No. 2.4e+02;  
Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;  
QY 1 ALNVLMA 7  
Db 2 SLNVLLA 8  
RESULT 17  
US-09-864-761-35481  
Sequence 35481, Application US/09864761  
Patent No. US20020048763A1  
GENERAL INFORMATION:  
APPLICANT: Penn, Sharron G.  
APPLICANT: Rank, David R.  
APPLICANT: Hanzel, David K.  
APPLICANT: Chen, Wensheng  
TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR  
TITLE OF INVENTION: GENE EXPRESSION ANALYSIS BY MICROARRAY  
FILE REFERENCE: Aecmica-X-1  
CURRENT APPLICATION NUMBER: US/09/864,761  
CURRENT FILING DATE: 2001-05-23  
PRIOR APPLICATION NUMBER: US 60/180,312  
PRIOR FILING DATE: 2000-02-04  
PRIOR APPLICATION NUMBER: US 60/207,456  
PRIOR FILING DATE: 2000-05-26  
PRIOR APPLICATION NUMBER: US 09/632,366  
PRIOR FILING DATE: 2000-08-03  
PRIOR APPLICATION NUMBER: GB 24263.6  
PRIOR FILING DATE: 2000-10-04  
PRIOR APPLICATION NUMBER: US 60/236,359  
PRIOR FILING DATE: 2000-09-27  
PRIOR APPLICATION NUMBER: PCT/US01/00666  
PRIOR FILING DATE: 2001-01-30  
PRIOR APPLICATION NUMBER: PCT/US01/00667  
PRIOR FILING DATE: 2001-01-30  
PRIOR APPLICATION NUMBER: PCT/US01/00664  
PRIOR FILING DATE: 2001-01-30  
PRIOR APPLICATION NUMBER: PCT/US01/00669  
PRIOR FILING DATE: 2001-01-30  
PRIOR APPLICATION NUMBER: PCT/US01/00665  
PRIOR FILING DATE: 2001-01-30  
PRIOR APPLICATION NUMBER: PCT/US01/00668  
PRIOR FILING DATE: 2001-01-30  
PRIOR APPLICATION NUMBER: PCT/US01/00663  
PRIOR FILING DATE: 2001-01-30  
PRIOR APPLICATION NUMBER: PCT/US01/00662  
PRIOR FILING DATE: 2001-01-30  
PRIOR APPLICATION NUMBER: PCT/US01/00661  
PRIOR FILING DATE: 2001-01-30  
PRIOR APPLICATION NUMBER: PCT/US01/00670  
PRIOR FILING DATE: 2001-01-30  
PRIOR APPLICATION NUMBER: US 60/234,687  
PRIOR FILING DATE: 2000-09-21  
PRIOR APPLICATION NUMBER: US 09/608,408  
PRIOR FILING DATE: 2000-06-30  
PRIOR APPLICATION NUMBER: US 09/774,203  
PRIOR FILING DATE: 2001-01-29  
NUMBER OF SEQ ID NOS: 49117  
SOFTWARE: Annonax Sequence Listing Engine vers. 1.1  
SEQ ID NO 35481  
LENGTH: 27  
TYPE: PRT

ORGANISM: Homo sapiens  
FEATURE:  
OTHER INFORMATION: MAP TO AB020871.1  
OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 3.8  
OTHER INFORMATION: EXPRESSED IN HEL100, SIGNAL = 2.8  
OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 20  
OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 3.2  
OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 7.2  
OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 4.3  
OTHER INFORMATION: EXPRESSED IN BT474, SIGNAL = 2.8  
OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 9.8  
OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 3.6  
OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 12  
OTHER INFORMATION: EST\_HUMAN HIT: AA948008.1, EVALUE 2.00e-05  
US-09-864-761-35481  
Query Match 77.4%; Score 24; DB 9; Length 27;  
Best Local Similarity 83.3%; Pred. No. 1.2e+02;  
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;  
QY 2 LNLVMA 7  
Db 4 LNLVMA 9  
RESULT 18  
US-10-106-698-7834  
Sequence 7834, Application US/10106698  
Publication No. US20030109690A1  
GENERAL INFORMATION:  
APPLICANT: Ruben et al.  
TITLE OF INVENTION: Colon and Colon Cancer Associated Polynucleotides and Polypeptides  
FILE REFERENCE: PA005P1  
CURRENT APPLICATION NUMBER: US/10/106,698  
CURRENT FILING DATE: 2002-03-27  
PRIOR APPLICATION NUMBER: PCT/US00/26524  
PRIOR FILING DATE: 2000-09-28  
PRIOR APPLICATION NUMBER: US 60/157,137  
PRIOR FILING DATE: 1999-09-29  
PRIOR APPLICATION NUMBER: US 60/163,280  
PRIOR FILING DATE: 1999-11-03  
NUMBER OF SEQ ID NOS: 8564  
SOFTWARE: PatentIn Ver. 3.0  
SEQ ID NO 7834  
LENGTH: 32  
TYPE: PRT  
ORGANISM: Homo sapiens  
US-10-106-698-7834  
Query Match 77.4%; Score 24; DB 15; Length 32;  
Best Local Similarity 83.3%; Pred. No. 1.4e+02;  
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;  
QY 1 ALNVLMA 6  
Db 22 SLNVLM 27  
RESULT 19  
US-09-738-626-4982  
Sequence 4982, Application US/09738626  
Publication No. US20020197605A1  
GENERAL INFORMATION:  
APPLICANT: NAKAGAWA, SATOSHI  
APPLICANT: MIZOGUCHI, HIROSHI  
APPLICANT: ANDO, SEIKO  
APPLICANT: HAYASHI, MIKIRO  
APPLICANT: OCHIAI, KEIKO  
APPLICANT: YOKOI, HARUHIKO  
APPLICANT: TATEISHI, NAKO  
APPLICANT: SENO, AKIHIRO  
APPLICANT: IKEDA, MASATO  
APPLICANT: OZAKI, AKIO

```
; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
; FILE REFERENCE: 249-125
; CURRENT APPLICATION NUMBER: US/09/738,626
; CURRENT FILING DATE: 2000-12-18
; PRIOR APPLICATION NUMBER: JP 99/377484
; PRIOR FILING DATE: 1999-12-16
; PRIOR APPLICATION NUMBER: JP 00/159162
; PRIOR FILING DATE: 2000-04-07
; PRIOR APPLICATION NUMBER: JP 00/280988
; PRIOR FILING DATE: 2000-08-03
; NUMBER OF SEQ ID NOS: 7059
; SOFTWARE: PatentIn ver. 3.0
; SEQ ID NO 4982
; LENGTH: 64
; TYPE: PRT
; ORGANISM: Corynebacterium glutamicum
US-09-738-626-4982

Query Match          77.4%  Score 24;  DB 10;  Length 64;
Best Local Similarity 83.3%;  Pred. No. 2.9e+02;
Matches 5;  Conservative 1;  Mismatches 0;  Indels 0;  Gaps 0;

QY      1  ALNVLN 6
Db      24  ALNVL 29

RESULT 20
US-09-764-860-538
; Sequence 538, Application US/09764860
; Patent No. US20020094953A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PC008
; CURRENT APPLICATION NUMBER: US/09/764,860
; CURRENT FILING DATE: 2001-01-17
; Number of SEQ ID NOS: 1198
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 538
; LENGTH: 65
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (3)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
US-09-764-860-538

Query Match          77.4%  Score 24;  DB 9;  Length 65;
Best Local Similarity 66.7%;  Pred. No. 3e+02;
Matches 4;  Conservative 2;  Mismatches 0;  Indels 0;  Gaps 0;

QY      2  LNVLMA 7
Db      28  LNIIMA 33

RESULT 21
US-10-212-872-538
; Sequence 538, Application US/10212872
; Publication No. US20030215893A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PC008C2
; CURRENT APPLICATION NUMBER: US/10/212,872
; CURRENT FILING DATE: 2002-08-07
; Prior application removed - See File Wrapper or Palm
; NUMBER OF SEQ ID NOS: 1198
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 538

; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
; FILE REFERENCE: 249-125
; CURRENT APPLICATION NUMBER: US/09/738,626
; CURRENT FILING DATE: 2000-12-18
; PRIOR APPLICATION NUMBER: JP 99/377484
; PRIOR FILING DATE: 1999-12-16
; PRIOR APPLICATION NUMBER: JP 00/159162
; PRIOR FILING DATE: 2000-04-07
; PRIOR APPLICATION NUMBER: JP 00/280988
; PRIOR FILING DATE: 2000-08-03
; NUMBER OF SEQ ID NOS: 7059
; SOFTWARE: PatentIn ver. 3.0
; SEQ ID NO 4982
; LENGTH: 64
; TYPE: PRT
; ORGANISM: Corynebacterium glutamicum
US-09-738-626-4982

Query Match          77.4%  Score 24;  DB 12;  Length 65;
Best Local Similarity 66.7%;  Pred. No. 3e+02;
Matches 4;  Conservative 2;  Mismatches 0;  Indels 0;  Gaps 0;

QY      2  LNVLMA 7
Db      28  LNIIMA 33

RESULT 22
US-10-074-095-538
; Sequence 538, Application US/10074095
; Publication No. US2003007704A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PC008C1
; CURRENT APPLICATION NUMBER: US/10/074,095
; CURRENT FILING DATE: 2002-02-14
; PRIOR APPLICATION NUMBER: 09/764,860
; PRIOR FILING DATE: 2001-01-17
; PRIOR APPLICATION NUMBER: 60/179,065
; PRIOR FILING DATE: 2000-01-31
; PRIOR APPLICATION NUMBER: 60/180,628
; PRIOR FILING DATE: 2000-02-04
; PRIOR APPLICATION NUMBER: 60/214,886
; PRIOR FILING DATE: 2000-06-28
; PRIOR APPLICATION NUMBER: 60/217,487
; PRIOR FILING DATE: 2000-07-11
; PRIOR APPLICATION NUMBER: 60/225,758
; PRIOR FILING DATE: 2000-08-14
; PRIOR APPLICATION NUMBER: 60/220,963
; PRIOR FILING DATE: 2000-07-26
; PRIOR APPLICATION NUMBER: 60/217,496
; PRIOR FILING DATE: 2000-07-11
; PRIOR APPLICATION NUMBER: 60/225,447
; PRIOR FILING DATE: 2000-08-14
; PRIOR APPLICATION NUMBER: 60/218,290
; PRIOR FILING DATE: 2000-07-14
; PRIOR APPLICATION NUMBER: 60/225,757
; PRIOR FILING DATE: 2000-08-14
; PRIOR APPLICATION NUMBER: 60/226,868
; PRIOR FILING DATE: 2000-08-22
; PRIOR APPLICATION NUMBER: 60/216,647
; PRIOR FILING DATE: 2000-07-07
; PRIOR APPLICATION NUMBER: 60/225,267
; PRIOR FILING DATE: 2000-08-14
; PRIOR APPLICATION NUMBER: 60/216,880
; PRIOR FILING DATE: 2000-07-07
; PRIOR APPLICATION NUMBER: 60/225,270
; PRIOR FILING DATE: 2000-08-14
; PRIOR APPLICATION NUMBER: 60/251,869
; PRIOR FILING DATE: 2000-12-08
; PRIOR APPLICATION NUMBER: 60/235,834
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: 60/234,274
; PRIOR FILING DATE: 2000-09-21
; PRIOR APPLICATION NUMBER: 60/234,223
; PRIOR FILING DATE: 2000-09-21
; PRIOR APPLICATION NUMBER: 60/228,924
; PRIOR FILING DATE: 2000-08-30
; PRIOR APPLICATION NUMBER: 60/224,518
; PRIOR FILING DATE: 2000-08-14
; PRIOR APPLICATION NUMBER: 60/236,369
```

100



```
; PRIOR FILING DATE: 2000-09-08
Query Match      77.4%; Score 24; DB 15; Length 65;
Best Local Similarity 66.7%; Pred. No. 3e+02;
Matches 4; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Qy  2 LNVLMA 7
Db  28 LNIIMA 33

RESULT 23
US-09-925-300-1456
; Sequence 1456, Application US/09925300
; Patent No. US20020151681A1
; GENERAL INFORMATION:
; APPLICANT: Craig Rosen,
; APPLICANT: Steve Ruben
; TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies
; FILE REFERENCE: PA101
; CURRENT APPLICATION NUMBER: US/09/925,300
; CURRENT FILING DATE: 2001-08-10
; PRIOR APPLICATION NUMBER: PCT/US00/05988
; PRIOR FILING DATE: 2000-03-08
; PRIOR APPLICATION NUMBER: 60/124,270
; PRIOR FILING DATE: 1999-03-12
; NUMBER OF SEQ ID NOS: 1890
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1456
; LENGTH: 71
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (10)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
US-09-925-300-1456

Query Match      77.4%; Score 24; DB 10; Length 71;
Best Local Similarity 66.7%; Pred. No. 3.3e+02;
Matches 4; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Qy  2 LNVLMA 7
Db  35 LNIIMA 40

RESULT 24
US-10-029-386-29049
; Sequence 29049, Application US/10029386
; Publication No. US20030194704A1
; GENERAL INFORMATION:
; APPLICANT: Penn, Sharron G.
; APPLICANT: Rank, David R.
; APPLICANT: Hanzel, David K.
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
; TITLE OF INVENTION: EXPRESSION ANALYSIS TWO
; FILE REFERENCE: AEOMICA-X-2
; CURRENT APPLICATION NUMBER: US/10/029,386
; CURRENT FILING DATE: 2001-12-20
; NUMBER OF SEQ ID NOS: 34288
; SOFTWARE: Annonax Sequence Listing Engine vers. 1.1
; SEQ ID NO 29049
; LENGTH: 96
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: MAP TO CHR19.1
; OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 1.5
; OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 1.3
; OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 1.2
; OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 1.3
; OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 1.5
```

```
; OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 1.8
US-10-029-386-29049

Query Match      77.4%; Score 24; DB 12; Length 96;
Best Local Similarity 83.3%; Pred. No. 4.5e+02;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy  2 LNVLMA 7
Db  13 VNVLMA 18

RESULT 25
US-09-900-147-2
; Sequence 2, Application US/09900147
; Patent No. US20020103121A1
; GENERAL INFORMATION:
; APPLICANT: La Thangue, Nicholas B
; APPLICANT: Bandara, Lasantha R
; TITLE OF INVENTION: Peptide antagonists of DP transcription factors
; FILE REFERENCE: 620-67
; CURRENT APPLICATION NUMBER: US/09/900,147
; CURRENT FILING DATE: 2001-07-09
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 09/308,935
; PRIOR FILING DATE: EARLIER FILING DATE: 1999-05-27
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: GB 9626589.7
; PRIOR FILING DATE: EARLIER FILING DATE: 1996-12-20
; NUMBER OF SEQ ID NOS: 18
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2
; LENGTH: 9
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Synthetic
US-09-900-147-2

Query Match      74.2%; Score 23; DB 10; Length 9;
Best Local Similarity 100.0%; Pred. No. 7e+05;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy  3 NVLMA 7
Db  1 NVLMA 5

RESULT 26
US-09-864-761-41818
; Sequence 41818, Application US/09864761
; Patent No. US20020048763A1
; GENERAL INFORMATION:
; APPLICANT: Penn, Sharron G.
; APPLICANT: Rank, David R.
; APPLICANT: Hanzel, David K.
; APPLICANT: Chen, Wensheng
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
; TITLE OF INVENTION: GENE EXPRESSION ANALYSIS BY MICROARRAY
; FILE REFERENCE: Aeomica-X-1
; CURRENT APPLICATION NUMBER: US/09/864,761
; CURRENT FILING DATE: 2001-05-23
; PRIOR APPLICATION NUMBER: US 60/180,312
; PRIOR FILING DATE: 2000-02-04
; PRIOR APPLICATION NUMBER: US 60/207,456
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: US 09/632,366
; PRIOR FILING DATE: 2000-08-03
; PRIOR APPLICATION NUMBER: GB 24263.6
; PRIOR FILING DATE: 2000-10-04
; PRIOR APPLICATION NUMBER: US 60/236,359
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: PCT/US01/00666
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00667
```

```

1. PRIOR FILING DATE: 2001-01-30
2. PRIOR APPLICATION NUMBER: PCT/US01/00664
3. PRIOR FILING DATE: 2001-01-30
4. PRIOR APPLICATION NUMBER: PCT/US01/00669
5. PRIOR FILING DATE: 2001-01-30
6. PRIOR APPLICATION NUMBER: PCT/US01/00665
7. PRIOR FILING DATE: 2001-01-30
8. PRIOR APPLICATION NUMBER: PCT/US01/00668
9. PRIOR FILING DATE: 2001-01-30
10. PRIOR APPLICATION NUMBER: PCT/US01/00663
11. PRIOR FILING DATE: 2001-01-30
12. PRIOR APPLICATION NUMBER: PCT/US01/00662
13. PRIOR FILING DATE: 2001-01-30
14. PRIOR APPLICATION NUMBER: PCT/US01/00661
15. PRIOR FILING DATE: 2001-01-30
16. PRIOR APPLICATION NUMBER: PCT/US01/00670
17. PRIOR FILING DATE: 2001-01-30
18. PRIOR APPLICATION NUMBER: US 60/234,687
19. PRIOR FILING DATE: 2000-09-21
20. PRIOR APPLICATION NUMBER: US 09/608,408
21. PRIOR FILING DATE: 2000-06-30
22. PRIOR APPLICATION NUMBER: US 09/774,203
23. PRIOR FILING DATE: 2001-01-29
24. NUMBER OF SEQ ID NOS: 49117
25. SOFTWARE: Annonmax Sequence Listing Engine vers. 1.1
26. SEQ ID NO 41818
27. LENGTH: 20
28. TYPE: PRT
29. ORGANISM: Homo sapiens
30. FEATURE:
31. OTHER INFORMATION: MAP TO AC025289.1
32. OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 1.5
33. OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 2
34. OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 1.9
35. OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 1.6
36. OTHER INFORMATION: EST_HUMAN HIT: AW131202.1, EVALUAE 5.00e-04
37. US-09-864-761-41818

Query Match 74.2%; Score 23; DB 9; Length 20;
Best Local Similarity 66.7%; Pred. No. 1.4e+02;
Matches 4; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 2 LNVLMA 7
DB 4 LNILIA 9

RESULT 27
US-09-900-147-4
; Sequence 4, Application US/09900147
; Patent No. US20020103121A1
; GENERAL INFORMATION:
; APPLICANT: La Thangue, Nicholas B
; APPLICANT: Bandara, Lasantha R
; TITLE OF INVENTION: Peptide antagonists of DP transcription factors
; FILE REFERENCE: 620-67
; CURRENT APPLICATION NUMBER: US/09/900,147
; CURRENT FILING DATE: 2001-07-09
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 09/308,935
; PRIOR FILING DATE: EARLIER FILING DATE: 1999-05-27
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: GB 9626589.7
; PRIOR FILING DATE: EARLIER FILING DATE: 1996-12-20
; NUMBER OF SEQ ID NOS: 18
; SOFTWARE: PatentIn ver. 2.1
; SEQ ID NO 4
; LENGTH: 20
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Synthetic
; US-09-900-147-4

Query Match 74.2%; Score 23; DB 10; Length 20;

```

```
Query Match      74.2%; Score 23; DB 9; Length 26;
Best Local Similarity 83.3%; Pred. No. 1.9e+02;
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY      1 ALNVLM 6
      |||||
Db      10 ALNVFM 15

RESULT 29
US-10-024-652-27
; Sequence 27, Application US/10024652
; Publication No. US20030219738A1
; GENERAL INFORMATION:
; APPLICANT: Agensys, Inc
; APPLICANT: Challita-Eid, Pia M.
; APPLICANT: Paris, Mary
; APPLICANT: Afar, Daniel E.H.
; APPLICANT: Hubert, Rene S.
; APPLICANT: Mitchell, Steve Chappell
; APPLICANT: Levin, Elana
; APPLICANT: Morrison, Karen Jane Meyrick
; APPLICANT: Raitano, Arthur B.
; APPLICANT: Jakobovits, Aya
; TITLE OF INVENTION: Nucleic Acid and Encoded Zinc
; TITLE OF INVENTION: Transporter Protein Entitled 108PSH8 Useful in Treatment and
; TITLE OF INVENTION: Detection of Cancer
; FILE REFERENCE: 51158-20025.00
; CURRENT APPLICATION NUMBER: US/10/024,652
; CURRENT FILING DATE: 2002-06-28
; PRIOR FILING DATE: 2000-12-15
; NUMBER OF SEQ ID NOS: 2598
; SOFTWARE: Fast-SEQ for Windows Version 4.0
; SEQ ID NO 27
; LENGTH: 9
; TYPE: PRT
; ORGANISM: homo sapien
US-10-024-652-27

Query Match      71.0%; Score 22; DB 12; Length 9;
Best Local Similarity 66.7%; Pred. No. 7e+05;
Matches 4; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY      1 ALNVLM 6
      ||:|:|
Db      1 AVNVIM 6

RESULT 31
US-10-024-652-348
; Sequence 348, Application US/10024652
; Publication No. US20030219738A1
; GENERAL INFORMATION:
; APPLICANT: Agensys, Inc
; APPLICANT: Challita-Eid, Pia M.
; APPLICANT: Paris, Mary
; APPLICANT: Afar, Daniel E.H.
; APPLICANT: Hubert, Rene S.
; APPLICANT: Mitchell, Steve Chappell
; APPLICANT: Levin, Elana
; APPLICANT: Morrison, Karen Jane Meyrick
; APPLICANT: Raitano, Arthur B.
; APPLICANT: Jakobovits, Aya
; TITLE OF INVENTION: Nucleic Acid and Encoded Zinc
; TITLE OF INVENTION: Transporter Protein Entitled 108PSH8 Useful in Treatment and
; TITLE OF INVENTION: Detection of Cancer
; FILE REFERENCE: 51158-20025.00
; CURRENT APPLICATION NUMBER: US/10/024,652
; CURRENT FILING DATE: 2002-06-28
; PRIOR FILING DATE: 2000-12-15
; NUMBER OF SEQ ID NOS: 2598
; SOFTWARE: Fast-SEQ for Windows Version 4.0
; SEQ ID NO 348
; LENGTH: 9
; TYPE: PRT
; ORGANISM: homo sapien
US-10-024-652-348

Query Match      71.0%; Score 22; DB 12; Length 9;
Best Local Similarity 66.7%; Pred. No. 7e+05;
Matches 4; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY      1 ALNVLM 6
      ||:|:|
Db      1 AVNVIM 6

RESULT 32
US-10-024-652-412
; Sequence 412, Application US/10024652
; Publication No. US20030219738A1
; GENERAL INFORMATION:
; APPLICANT: Agensys, Inc
; APPLICANT: Challita-Eid, Pia M.
; APPLICANT: Paris, Mary
; APPLICANT: Afar, Daniel E.H.
; APPLICANT: Hubert, Rene S.
; APPLICANT: Mitchell, Steve Chappell
; APPLICANT: Levin, Elana
; APPLICANT: Morrison, Karen Jane Meyrick
; APPLICANT: Raitano, Arthur B.
; APPLICANT: Jakobovits, Aya
; TITLE OF INVENTION: Nucleic Acid and Encoded Zinc
; TITLE OF INVENTION: Transporter Protein Entitled 108PSH8 Useful in Treatment and
```

Tue Feb 17 11:55:56 2004

```
; TITLE OF INVENTION: Detection of Cancer
; FILE REFERENCE: 51158-20025.00
; CURRENT APPLICATION NUMBER: US/10/024,652
; CURRENT FILING DATE: 2002-06-28
; PRIOR APPLICATION NUMBER: 60/256,210
; PRIOR FILING DATE: 2000-12-15
; NUMBER OF SEQ ID NOS: 2598
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 412
; LENGTH: 9
; TYPE: PRT
; ORGANISM: homo sapien
US-10-024-652-412

Query Match          71.0%; Score 22; DB 12; Length 9;
Best Local Similarity 66.7%; Pred. No. 7e+05;
Matches 4; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ALNVLM 6
   |.|||:|
Db 1 AVNVIM 6

RESULT 33
US-10-024-652-444
; Sequence 444, Application US/10024652
; Publication No. US20030219738A1
; GENERAL INFORMATION:
; APPLICANT: Agensys, Inc
; APPLICANT: Challita-Eid, Pia M.
; APPLICANT: Faris, Mary
; APPLICANT: Afar, Daniel E.H.
; APPLICANT: Hubert, Rene S.
; APPLICANT: Mitchell, Steve Chappell
; APPLICANT: Levin, Elana
; APPLICANT: Morrison, Karen Jane Meyrick
; APPLICANT: Raitano, Arthur B.
; APPLICANT: Jakobovits, Aya
; TITLE OF INVENTION: Nucleic Acid and Encoded Zinc
; TITLE OF INVENTION: Transporter Protein Entitled 108PSH8 Useful in Treatment and
; TITLE OF INVENTION: Detection of Cancer
; FILE REFERENCE: 51158-20025.00
; CURRENT APPLICATION NUMBER: US/10/024,652
; CURRENT FILING DATE: 2002-06-28
; PRIOR APPLICATION NUMBER: 60/256,210
; PRIOR FILING DATE: 2000-12-15
; NUMBER OF SEQ ID NOS: 2598
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 444
; LENGTH: 9
; TYPE: PRT
; ORGANISM: homo sapien
US-10-024-652-444

Query Match          71.0%; Score 22; DB 12; Length 9;
Best Local Similarity 66.7%; Pred. No. 7e+05;
Matches 4; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ALNVLM 6
   |.|||:|
Db 2 AVNVIM 7

RESULT 34
US-10-024-652-501
; Sequence 501, Application US/10024652
; Publication No. US20030219738A1
; GENERAL INFORMATION:
; APPLICANT: Agensys, Inc
; APPLICANT: Challita-Eid, Pia M.
; APPLICANT: Faris, Mary
; APPLICANT: Afar, Daniel E.H.
; APPLICANT: Hubert, Rene S.
; APPLICANT: Mitchell, Steve Chappell
; APPLICANT: Levin, Elana
; APPLICANT: Morrison, Karen Jane Meyrick
; APPLICANT: Raitano, Arthur B.
; APPLICANT: Jakobovits, Aya
; TITLE OF INVENTION: Nucleic Acid and Encoded Zinc
; TITLE OF INVENTION: Transporter Protein Entitled 108PSH8 Useful in Treatment and
; TITLE OF INVENTION: Detection of Cancer
; FILE REFERENCE: 51158-20025.00
; CURRENT APPLICATION NUMBER: US/10/024,652
; CURRENT FILING DATE: 2002-06-28
; PRIOR APPLICATION NUMBER: 60/256,210
; PRIOR FILING DATE: 2000-12-15
; NUMBER OF SEQ ID NOS: 2598
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 501
; LENGTH: 9
; TYPE: PRT
; ORGANISM: homo sapien
US-10-024-652-501

Query Match          71.0%; Score 22; DB 12; Length 9;
Best Local Similarity 66.7%; Pred. No. 7e+05;
Matches 4; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ALNVLM 6
   |.|||:|
Db 2 AVNVIM 7

RESULT 35
US-10-024-652-626
; Sequence 626, Application US/10024652
; Publication No. US20030219738A1
; GENERAL INFORMATION:
; APPLICANT: Agensys, Inc
; APPLICANT: Challita-Eid, Pia M.
; APPLICANT: Faris, Mary
; APPLICANT: Afar, Daniel E.H.
; APPLICANT: Hubert, Rene S.
; APPLICANT: Mitchell, Steve Chappell
; APPLICANT: Levin, Elana
; APPLICANT: Morrison, Karen Jane Meyrick
; APPLICANT: Raitano, Arthur B.
; APPLICANT: Jakobovits, Aya
; TITLE OF INVENTION: Nucleic Acid and Encoded Zinc
; TITLE OF INVENTION: Transporter Protein Entitled 108PSH8 Useful in Treatment and
; TITLE OF INVENTION: Detection of Cancer
; FILE REFERENCE: 51158-20025.00
; CURRENT APPLICATION NUMBER: US/10/024,652
; CURRENT FILING DATE: 2002-06-28
; PRIOR APPLICATION NUMBER: 60/256,210
; PRIOR FILING DATE: 2000-12-15
; NUMBER OF SEQ ID NOS: 2598
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 626
; LENGTH: 9
; TYPE: PRT
; ORGANISM: homo sapien
US-10-024-652-626

Query Match          71.0%; Score 22; DB 12; Length 9;
Best Local Similarity 66.7%; Pred. No. 7e+05;
Matches 4; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ALNVLM 6
   |.|||:|
Db 2 AVNVIM 7

RESULT 36
US-10-024-652-636
; Sequence 636, Application US/10024652
```

```

; Publication No. US20030219738A1
; GENERAL INFORMATION:
; APPLICANT: Agensys, Inc
; APPLICANT: Challita-Eid, Pia M.
; APPLICANT: Faris, Mary
; APPLICANT: Afar, Daniel E.H.
; APPLICANT: Hubert, Rene S.
; APPLICANT: Mitchell, Steve Chappell
; APPLICANT: Levin, Elana
; APPLICANT: Morrison, Karen Jane Meyrick
; APPLICANT: Raitano, Arthur B.
; APPLICANT: Jakobovits, Aya
; TITLE OF INVENTION: Nucleic Acid and Encoded Zinc
; TITLE OF INVENTION: Transporter Protein Entitled 108PSH8 Useful in Treatment and
; TITLE OF INVENTION: Detection of Cancer
; FILE REFERENCE: 51158-20025.00
; CURRENT APPLICATION NUMBER: US/10/024,652
; CURRENT FILING DATE: 2002-06-28
; PRIOR FILING DATE: 2000-12-15
; NUMBER OF SEQ ID NOS: 2598
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 636
; LENGTH: 9
; TYPE: PRT
; ORGANISM: homo sapien
US-10-024-652-636

Query Match 71.0%; Score 22; DB 12; Length 9;
Best Local Similarity 66.7%; Pred. No. 7e+05;
Matches 4; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 ALNVLM 6
DB 4 AVNVIM 9

RESULT 37
US-10-024-652-926
; Sequence 926, Application US/10024652
; Publication No. US20030219738A1
; GENERAL INFORMATION:
; APPLICANT: Agensys, Inc
; APPLICANT: Challita-Eid, Pia M.
; APPLICANT: Faris, Mary
; APPLICANT: Afar, Daniel E.H.
; APPLICANT: Hubert, Rene S.
; APPLICANT: Mitchell, Steve Chappell
; APPLICANT: Levin, Elana
; APPLICANT: Morrison, Karen Jane Meyrick
; APPLICANT: Raitano, Arthur B.
; APPLICANT: Jakobovits, Aya
; TITLE OF INVENTION: Nucleic Acid and Encoded Zinc
; TITLE OF INVENTION: Transporter Protein Entitled 108PSH8 Useful in Treatment and
; TITLE OF INVENTION: Detection of Cancer
; FILE REFERENCE: 51158-20025.00
; CURRENT APPLICATION NUMBER: US/10/024,652
; CURRENT FILING DATE: 2002-06-28
; PRIOR FILING DATE: 2000-12-15
; NUMBER OF SEQ ID NOS: 2598
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 926
; LENGTH: 9
; TYPE: PRT
; ORGANISM: homo sapien
US-10-024-652-926

Query Match 71.0%; Score 22; DB 12; Length 9;
Best Local Similarity 66.7%; Pred. No. 7e+05;
Matches 4; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 ALNVLM 6
DB 4 AVNVIM 9

RESULT 38
US-10-024-652-976
; Sequence 976, Application US/10024652
; Publication No. US20030219738A1
; GENERAL INFORMATION:
; APPLICANT: Agensys, Inc
; APPLICANT: Challita-Eid, Pia M.
; APPLICANT: Faris, Mary
; APPLICANT: Afar, Daniel E.H.
; APPLICANT: Hubert, Rene S.
; APPLICANT: Mitchell, Steve Chappell
; APPLICANT: Levin, Elana
; APPLICANT: Morrison, Karen Jane Meyrick
; APPLICANT: Raitano, Arthur B.
; APPLICANT: Jakobovits, Aya
; TITLE OF INVENTION: Nucleic Acid and Encoded Zinc
; TITLE OF INVENTION: Transporter Protein Entitled 108PSH8 Useful in Treatment and
; TITLE OF INVENTION: Detection of Cancer
; FILE REFERENCE: 51158-20025.00
; CURRENT APPLICATION NUMBER: US/10/024,652
; CURRENT FILING DATE: 2002-06-28
; PRIOR FILING DATE: 2000-12-15
; NUMBER OF SEQ ID NOS: 2598
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 976
; LENGTH: 9
; TYPE: PRT
; ORGANISM: homo sapien
US-10-024-652-976

Query Match 71.0%; Score 22; DB 12; Length 9;
Best Local Similarity 66.7%; Pred. No. 7e+05;
Matches 4; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 ALNVLM 6
DB 3 AVNVIM 8

RESULT 39
US-10-024-652-977
; Sequence 977, Application US/10024652
; Publication No. US20030219738A1
; GENERAL INFORMATION:
; APPLICANT: Agensys, Inc
; APPLICANT: Challita-Eid, Pia M.
; APPLICANT: Faris, Mary
; APPLICANT: Afar, Daniel E.H.
; APPLICANT: Hubert, Rene S.
; APPLICANT: Mitchell, Steve Chappell
; APPLICANT: Levin, Elana
; APPLICANT: Morrison, Karen Jane Meyrick
; APPLICANT: Raitano, Arthur B.
; APPLICANT: Jakobovits, Aya
; TITLE OF INVENTION: Nucleic Acid and Encoded Zinc
; TITLE OF INVENTION: Transporter Protein Entitled 108PSH8 Useful in Treatment and
; TITLE OF INVENTION: Detection of Cancer
; FILE REFERENCE: 51158-20025.00
; CURRENT APPLICATION NUMBER: US/10/024,652
; CURRENT FILING DATE: 2002-06-28
; PRIOR FILING DATE: 2000-12-15
; NUMBER OF SEQ ID NOS: 2598
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 977
; LENGTH: 9
; TYPE: PRT
; ORGANISM: homo sapien
US-10-024-652-977
```

US-10-024-652-977

Query Match 71.0%; Score 22; DB 12; Length 9;  
Best Local Similarity 66.7%; Pred. No. 7e+05; Indels 0;  
Matches 4; Conservative 2; Mismatches 0; Gaps 0;

Qy 1 ALNVLM 6  
|:|:|:  
Db 2 AVNVIM 7

## RESULT 40

US-10-024-652-1076  
Sequence 1076, Application US/10024652  
Publication No. US20030219738A1

## GENERAL INFORMATION:

APPLICANT: Agensys, Inc  
APPLICANT: Challita-Eid, Pia M.  
APPLICANT: Faris, Mary  
APPLICANT: Afar, Daniel E.H.  
APPLICANT: Hubert, Rene S.  
APPLICANT: Mitchell, Steve Chappell  
APPLICANT: Levin, Elana  
APPLICANT: Morrison, Karen Jane Meyrick  
APPLICANT: Raitano, Arthur B.  
APPLICANT: Jakobovits, Aya  
TITLE OF INVENTION: Nucleic Acid and Encoded Zinc  
TITLE OF INVENTION: Transporter Protein Entitled 108PSH8 Useful in Treatment and  
TITLE OF INVENTION: Detection of Cancer  
FILE REFERENCE: 51158-20025.00  
CURRENT APPLICATION NUMBER: US/10/024.652  
CURRENT FILING DATE: 2002-06-28  
PRIOR FILING DATE: 2000-12-15  
NUMBER OF SEQ ID NOS: 2598  
SOFTWARE: FastSEQ for Windows Version 4.0  
SEQ ID NO 1076  
LENGTH: 9  
TYPE: PRT  
ORGANISM: homo sapien

US-10-024-652-1076

Query Match 71.0%; Score 22; DB 12; Length 9;  
Best Local Similarity 66.7%; Pred. No. 7e+05; Indels 0;  
Matches 4; Conservative 2; Mismatches 0; Gaps 0;

Qy 1 ALNVLM 6  
|:|:|:  
Db 1 AVNVIM 6

## RESULT 41

US-10-024-652-1219  
Sequence 1219, Application US/10024652  
Publication No. US20030219738A1

## GENERAL INFORMATION:

APPLICANT: Agensys, Inc  
APPLICANT: Challita-Eid, Pia M.  
APPLICANT: Faris, Mary  
APPLICANT: Afar, Daniel E.H.  
APPLICANT: Hubert, Rene S.  
APPLICANT: Mitchell, Steve Chappell  
APPLICANT: Levin, Elana  
APPLICANT: Morrison, Karen Jane Meyrick  
APPLICANT: Raitano, Arthur B.  
APPLICANT: Jakobovits, Aya  
TITLE OF INVENTION: Nucleic Acid and Encoded Zinc  
TITLE OF INVENTION: Transporter Protein Entitled 108PSH8 Useful in Treatment and  
TITLE OF INVENTION: Detection of Cancer  
FILE REFERENCE: 51158-20025.00  
CURRENT APPLICATION NUMBER: US/10/024.652  
CURRENT FILING DATE: 2002-06-28  
PRIOR APPLICATION NUMBER: 60/256,210

PRIOR FILING DATE: 2000-12-15  
NUMBER OF SEQ ID NOS: 2598  
SOFTWARE: FastSEQ for Windows Version 4.0  
SEQ ID NO 1219  
LENGTH: 9  
TYPE: PRT  
ORGANISM: homo sapien  
US-10-024-652-1219

Query Match 71.0%; Score 22; DB 12; Length 9;  
Best Local Similarity 66.7%; Pred. No. 7e+05; Indels 0;  
Matches 4; Conservative 2; Mismatches 0; Gaps 0;

Qy 1 ALNVLM 6  
|:|:|:  
Db 1 AVNVIM 6

## RESULT 42

US-10-024-652-1272  
Sequence 1272, Application US/10024652  
Publication No. US20030219738A1

## GENERAL INFORMATION:

APPLICANT: Agensys, Inc  
APPLICANT: Challita-Eid, Pia M.  
APPLICANT: Faris, Mary  
APPLICANT: Afar, Daniel E.H.  
APPLICANT: Hubert, Rene S.  
APPLICANT: Mitchell, Steve Chappell  
APPLICANT: Levin, Elana  
APPLICANT: Morrison, Karen Jane Meyrick  
APPLICANT: Raitano, Arthur B.  
APPLICANT: Jakobovits, Aya  
TITLE OF INVENTION: Nucleic Acid and Encoded Zinc  
TITLE OF INVENTION: Transporter Protein Entitled 108PSH8 Useful in Treatment and  
TITLE OF INVENTION: Detection of Cancer  
FILE REFERENCE: 51158-20025.00  
CURRENT APPLICATION NUMBER: US/10/024.652  
CURRENT FILING DATE: 2002-06-28  
PRIOR APPLICATION NUMBER: 60/256,210  
PRIOR FILING DATE: 2000-12-15  
NUMBER OF SEQ ID NOS: 2598  
SOFTWARE: FastSEQ for Windows Version 4.0  
SEQ ID NO 1272  
LENGTH: 9  
TYPE: PRT  
ORGANISM: homo sapien  
US-10-024-652-1272

Query Match 71.0%; Score 22; DB 12; Length 9;  
Best Local Similarity 66.7%; Pred. No. 7e+05; Indels 0;  
Matches 4; Conservative 2; Mismatches 0; Gaps 0;

Qy 1 ALNVLM 6  
|:|:|:  
Db 4 AVNVIM 9

## RESULT 43

US-10-024-652-1279  
Sequence 1279, Application US/10024652  
Publication No. US20030219738A1

## GENERAL INFORMATION:

APPLICANT: Agensys, Inc  
APPLICANT: Challita-Eid, Pia M.  
APPLICANT: Faris, Mary  
APPLICANT: Afar, Daniel E.H.  
APPLICANT: Hubert, Rene S.  
APPLICANT: Mitchell, Steve Chappell  
APPLICANT: Levin, Elana  
APPLICANT: Morrison, Karen Jane Meyrick  
APPLICANT: Raitano, Arthur B.  
APPLICANT: Jakobovits, Aya

; TITLE OF INVENTION: Nucleic Acid and Encoded Zinc  
; TITLE OF INVENTION: Transporter Protein Entitled 108P5H8 Useful in Treatment and  
; TITLE OF INVENTION: Detection of Cancer  
; FILE REFERENCE: 51158-20025.00  
; CURRENT APPLICATION NUMBER: US/10/024,652  
; CURRENT FILING DATE: 2002-06-28  
; PRIOR APPLICATION NUMBER: 60/256,210  
; PRIOR FILING DATE: 2000-12-15  
; NUMBER OF SEQ ID NOS: 2598  
; SOFTWARE: FastSEQ for Windows Version 4.0  
; SEQ ID NO 1279  
; LENGTH: 9  
; TYPE: PRT  
; ORGANISM: homo sapien  
US-10-024-652-1279

Query Match 71.0%; Score 22; DB 12; Length 9;  
Best Local Similarity 66.7%; Pred. No. 7e+05;  
Matches 4; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 ALNVLM 6  
|:|:|:  
Db 2 AVNVIM 7

## RESULT 44

US-10-024-652-1311  
; Sequence 1311, Application US/10024652  
; Publication No. US20030219738A1  
; GENERAL INFORMATION:  
; APPLICANT: Agensys, Inc  
; APPLICANT: Challita-Eid, Pia M.  
; APPLICANT: Faris, Mary  
; APPLICANT: Afar, Daniel E.H.  
; APPLICANT: Hubert, Rene S.  
; APPLICANT: Mitchell, Steve Chappell  
; APPLICANT: Levin, Elana  
; APPLICANT: Morrison, Karen Jane Meyrick  
; APPLICANT: Raitano, Arthur B.  
; APPLICANT: Jakobovits, Aya  
; TITLE OF INVENTION: Nucleic Acid and Encoded Zinc  
; TITLE OF INVENTION: Transporter Protein Entitled 108P5H8 Useful in Treatment and  
; FILE REFERENCE: 51158-20025.00  
; CURRENT APPLICATION NUMBER: US/10/024,652  
; CURRENT FILING DATE: 2002-06-28  
; PRIOR APPLICATION NUMBER: 60/256,210  
; PRIOR FILING DATE: 2000-12-15  
; NUMBER OF SEQ ID NOS: 2598  
; SOFTWARE: FastSEQ for Windows Version 4.0  
; SEQ ID NO 1311  
; LENGTH: 9  
; TYPE: PRT  
; ORGANISM: homo sapien  
US-10-024-652-1311

Query Match 71.0%; Score 22; DB 12; Length 9;  
Best Local Similarity 66.7%; Pred. No. 7e+05;  
Matches 4; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 ALNVLM 6  
|:|:|:  
Db 2 AVNVIM 7

## RESULT 45

US-10-024-652-1383  
; Sequence 1383, Application US/10024652  
; Publication No. US20030219738A1  
; GENERAL INFORMATION:  
; APPLICANT: Agensys, Inc  
; APPLICANT: Challita-Eid, Pia M.  
; APPLICANT: Faris, Mary

; APPLICANT: Afar, Daniel E.H.  
; APPLICANT: Hubert, Rene S.  
; APPLICANT: Mitchell, Steve Chappell  
; APPLICANT: Levin, Elana  
; APPLICANT: Morrison, Karen Jane Meyrick  
; APPLICANT: Raitano, Arthur B.  
; APPLICANT: Jakobovits, Aya  
; TITLE OF INVENTION: Nucleic Acid and Encoded Zinc  
; TITLE OF INVENTION: Transporter Protein Entitled 108P5H8 Useful in Treatment and  
; TITLE OF INVENTION: Detection of Cancer  
; FILE REFERENCE: 51158-20025.00  
; CURRENT APPLICATION NUMBER: US/10/024,652  
; CURRENT FILING DATE: 2002-06-28  
; PRIOR APPLICATION NUMBER: 60/256,210  
; PRIOR FILING DATE: 2000-12-15  
; NUMBER OF SEQ ID NOS: 2598  
; SOFTWARE: FastSEQ for Windows Version 4.0  
; SEQ ID NO 1383  
; LENGTH: 9  
; TYPE: PRT  
; ORGANISM: homo sapien  
US-10-024-652-1383

Query Match 71.0%; Score 22; DB 12; Length 9;  
Best Local Similarity 66.7%; Pred. No. 7e+05;  
Matches 4; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 ALNVLM 6  
|:|:|:  
Db 1 AVNVIM 6

Search completed: February 17, 2004, 11:02:23  
Job time : 11.2574 secs

**THIS PAGE BLANK (USPTO)**



GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: February 17, 2004, 10:50:13 ; Search time 5.05941 Seconds  
(without alignments)  
58.540 Million cell updates/sec

Title: US-09-900-147-7  
Perfect score: 31  
Sequence: 1 ALNVLMA 7

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 328717 seqs, 42310858 residues

Total number of hits satisfying chosen parameters: 228043

Minimum DB seq length: 0  
Maximum DB seq length: 100

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Issued Patents AA: \*  
1: /cgn2\_6/ptodata/1/iaa/5A COMB.pep.\*  
2: /cgn2\_6/ptodata/1/iaa/5B COMB.pep.\*  
3: /cgn2\_6/ptodata/1/iaa/6A COMB.pep.\*  
4: /cgn2\_6/ptodata/1/iaa/6B COMB.pep.\*  
5: /cgn2\_6/ptodata/1/iaa/PCTUS COMB.pep.\*  
6: /cgn2\_6/ptodata/1/iaa/backfiles1.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	31	100.0	7	3	US-09-308-935-7
2	31	100.0	14	3	US-09-308-935-11
3	31	100.0	16	3	US-09-308-935-5
4	31	100.0	17	2	US-08-428-131-13
5	31	100.0	17	3	US-09-078-596-13
6	31	100.0	19	3	US-09-308-935-3
7	31	100.0	19	3	US-09-308-935-16
8	31	100.0	30	3	US-09-308-935-6
9	31	100.0	37	3	US-09-308-935-1
10	31	100.0	72	2	US-08-428-131-11
11	31	100.0	72	3	US-09-078-596-11
12	31	100.0	74	3	US-08-894-139-10
13	27	87.1	19	3	US-09-308-935-15
14	25	80.6	87	4	US-09-152-060-89
15	24	77.4	25	1	US-08-484-493-5
16	24	77.4	25	2	US-08-484-494-5
17	24	77.4	25	2	US-08-345-212-5
18	24	77.4	25	3	US-09-249-003-5
19	24	77.4	25	4	US-09-685-844-5
20	24	77.4	63	1	US-08-194-338-14
21	23	74.2	9	3	US-09-308-935-2
22	23	74.2	20	3	US-09-308-935-4
23	23	74.2	22	1	US-08-004-139B-33
24	23	74.2	22	2	US-08-811-492-33
25	23	74.2	22	5	PCT-US96-10545A-33
26	23	74.2	40	3	US-09-215-221-18
27	23	74.2	67	4	US-09-134-001C-4216

28	22	71.0	11	3	US-09-308-935-9	Sequence 9, Appli
29	22	71.0	13	3	US-08-725-459B-11	Sequence 11, Appl
30	22	71.0	21	3	US-09-399-494-19	Sequence 19, Appl
31	22	71.0	22	3	US-08-725-459B-13	Sequence 13, Appl
32	22	71.0	24	3	US-08-725-459B-75	Sequence 75, Appl
33	22	71.0	29	1	US-08-524-757-16	Sequence 16, Appl
34	22	71.0	43	4	US-09-217-293-8	Sequence 8, Appli
35	22	71.0	52	3	US-08-725-459B-76	Sequence 76, Appl
36	22	71.0	55	2	US-08-152-721B-22	Sequence 22, Appl
37	22	71.0	56	3	US-08-725-459B-50	Sequence 50, Appl
38	22	71.0	63	4	US-09-328-352-5729	Sequence 5729, Ap
39	22	71.0	67	1	US-08-530-010-18	Sequence 18, Appl
40	22	71.0	67	2	US-08-484-101B-18	Sequence 18, Appl
41	22	71.0	67	3	US-08-714-524D-18	Sequence 18, Appl
42	22	71.0	84	3	US-08-725-459B-49	Sequence 49, Appl
43	22	71.0	90	3	US-08-725-459B-48	Sequence 48, Appl
44	22	71.0	94	3	US-08-725-459B-47	Sequence 47, Appl
45	22	71.0	98	3	US-08-725-459B-46	Sequence 46, Appl

ALIGNMENTS

RESULT 1  
US-09-308-935-7  
; Sequence 7, Application US/09308935  
; Patent No. 6268334  
; GENERAL INFORMATION:  
; APPLICANT: La Thangue, Nicholas B  
; APPLICANT: Bandara, Lasantha R  
; TITLE OF INVENTION: Peptide antagonists of DP transcription factors  
; FILE REFERENCE: 620-67  
; CURRENT APPLICATION NUMBER: US/09/308,935  
; CURRENT FILING DATE: 1999-05-27  
; EARLIER APPLICATION NUMBER: PCT/GB97/03506  
; EARLIER FILING DATE: 1997-12-22  
; EARLIER APPLICATION NUMBER: GB 9626589.7  
; EARLIER FILING DATE: 1996-12-20  
; NUMBER OF SEQ ID NOS: 18  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 7  
; LENGTH: 7  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Description of Artificial Sequence: Synthetic  
US-09-308-935-7

Query Match 100.0%; Score 31; DB 3; Length 7;  
Best Local Similarity 100.0%; Pred. No. 2.5e+05;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ALNVLMA 7  
| | | | |  
Db 1 ALNVLMA 7

RESULT 2  
US-09-308-935-11  
; Sequence 11, Application US/09308935  
; Patent No. 6268334  
; GENERAL INFORMATION:  
; APPLICANT: La Thangue, Nicholas B  
; APPLICANT: Bandara, Lasantha R  
; TITLE OF INVENTION: Peptide antagonists of DP transcription factors  
; FILE REFERENCE: 620-67  
; CURRENT APPLICATION NUMBER: US/09/308,935  
; CURRENT FILING DATE: 1999-05-27  
; EARLIER APPLICATION NUMBER: PCT/GB97/03506  
; EARLIER FILING DATE: 1997-12-22  
; EARLIER APPLICATION NUMBER: GB 9626589.7  
; EARLIER FILING DATE: 1996-12-20  
; NUMBER OF SEQ ID NOS: 18

```

; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 11
; LENGTH: 14
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Synthetic
US-09-308-935-11
Query Match      100.0%; Score 31; DB 3; Length 14;
Best Local Similarity 100.0%; Pred. No. 0.51; Indels 0; Gaps 0;
Matches 7; Conservative 0; Mismatches 0;

Qy 1 ALNVLMA 7
Db 6 ALNVLMA 12

RESULT 3
US-09-308-935-5
; Sequence 5, Application US/09308935
; Patent No. 6268334
; GENERAL INFORMATION:
; APPLICANT: La Thangue, Nicholas B
; APPLICANT: Bandara, Laseantha R
; TITLE OF INVENTION: Peptide antagonists of DP transcription factors
; FILE REFERENCE: 620-67
; CURRENT APPLICATION NUMBER: US/09/308,935
; CURRENT FILING DATE: 1999-05-27
; EARLIER APPLICATION NUMBER: PCT/GB97/03506
; EARLIER FILING DATE: 1997-12-22
; EARLIER APPLICATION NUMBER: GB 9626589.7
; EARLIER FILING DATE: 1996-12-20
; NUMBER OF SEQ ID NOS: 18
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 5
; LENGTH: 16
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Synthetic
US-09-308-935-5
Query Match      100.0%; Score 31; DB 3; Length 16;
Best Local Similarity 100.0%; Pred. No. 0.56; Indels 0; Gaps 0;
Matches 7; Conservative 0; Mismatches 0;

Qy 1 ALNVLMA 7
Db 5 ALNVLMA 11

RESULT 4
US-08-428-131-13
; Sequence 13, Application US/08428131
; Patent No. 5863757
; GENERAL INFORMATION:
; APPLICANT: La Thangue, Nicholas Barrie
; TITLE OF INVENTION: Transcription Factor DP-1
; NUMBER OF SEQUENCES: 14
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Nixon & Vanderhye
; STREET: 1100 No. 5863757th Glebe Road, 8th Floor
; CITY: Arlington
; STATE: Virginia
; COUNTRY: U.S.A.
; ZIP: 22201-4714
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: IBM PC compatible
; SOFTWARE: PatentIn Release #1.0, Version #1.25 (EPO)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/078,596
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/428,131
; FILING DATE: 23-JUN-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Arthur R. Crawford
; REGISTRATION NUMBER: 25,327
; REFERENCE/DOCKET NUMBER: 117-181
; TELEPHONE: (703) 816-4000
; TELEFAX: (703) 816-4100
; INFORMATION FOR SEQ ID NO: 13:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 17 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; MISMATCH TYPE: protein
US-09-078-596-13
Query Match      100.0%; Score 31; DB 3; Length 17;
Best Local Similarity 100.0%; Pred. No. 0.62; Indels 0; Gaps 0;
Matches 7; Conservative 0; Mismatches 0;

Qy 1 ALNVLMA 7
Db 6 ALNVLMA 12

RESULT 5
US-09-078-596-13
; Sequence 13, Application US/09078596
; Patent No. 6150116
; GENERAL INFORMATION:
; APPLICANT: La Thangue, Nicholas Barrie
; TITLE OF INVENTION: Transcription Factor DP-1
; NUMBER OF SEQUENCES: 14
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Nixon & Vanderhye
; STREET: 1100 No. 6150116th Glebe Road, 8th Floor
; CITY: Arlington
; STATE: Virginia
; COUNTRY: U.S.A.
; ZIP: 22201-4714
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: IBM PC compatible
; SOFTWARE: PatentIn Release #1.0, Version #1.25 (EPO)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/078,596
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/428,131
; FILING DATE: 23-JUN-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Arthur R. Crawford
; REGISTRATION NUMBER: 25,327
; REFERENCE/DOCKET NUMBER: 117-181
; TELEPHONE: (703) 816-4000
; TELEFAX: (703) 816-4100
; INFORMATION FOR SEQ ID NO: 13:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 17 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; MISMATCH TYPE: protein
US-09-078-596-13
Query Match      100.0%; Score 31; DB 3; Length 17;
Best Local Similarity 100.0%; Pred. No. 0.62; Indels 0; Gaps 0;
Matches 7; Conservative 0; Mismatches 0;

Qy 1 ALNVLMA 7
Db 6 ALNVLMA 12

RESULT 5
US-09-078-596-13
; Sequence 13, Application US/09078596
; Patent No. 6150116
; GENERAL INFORMATION:
; APPLICANT: La Thangue, Nicholas Barrie
; TITLE OF INVENTION: Transcription Factor DP-1
; NUMBER OF SEQUENCES: 14
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Nixon & Vanderhye
; STREET: 1100 No. 6150116th Glebe Road, 8th Floor
; CITY: Arlington
; STATE: Virginia
; COUNTRY: U.S.A.
; ZIP: 22201-4714
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: IBM PC compatible
; SOFTWARE: PatentIn Release #1.0, Version #1.25 (EPO)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/078,596
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/428,131
; FILING DATE: 23-JUN-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Arthur R. Crawford
; REGISTRATION NUMBER: 25,327
; REFERENCE/DOCKET NUMBER: 117-181
; TELEPHONE: (703) 816-4000
; TELEFAX: (703) 816-4100
; INFORMATION FOR SEQ ID NO: 13:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 17 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; MISMATCH TYPE: protein
US-09-078-596-13
Query Match      100.0%; Score 31; DB 3; Length 17;
Best Local Similarity 100.0%; Pred. No. 0.62; Indels 0; Gaps 0;
Matches 7; Conservative 0; Mismatches 0;

Qy 1 ALNVLMA 7
Db 6 ALNVLMA 12

RESULT 5
US-09-078-596-13
; Sequence 13, Application US/09078596
; Patent No. 6150116
; GENERAL INFORMATION:
; APPLICANT: La Thangue, Nicholas Barrie
; TITLE OF INVENTION: Transcription Factor DP-1
; NUMBER OF SEQUENCES: 14
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Nixon & Vanderhye
; STREET: 1100 No. 6150116th Glebe Road, 8th Floor
; CITY: Arlington
; STATE: Virginia
; COUNTRY: U.S.A.
; ZIP: 22201-4714
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: IBM PC compatible
; SOFTWARE: PatentIn Release #1.0, Version #1.25 (EPO)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/078,596
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/428,131
; FILING DATE: 23-JUN-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Arthur R. Crawford
; REGISTRATION NUMBER: 25,327
; REFERENCE/DOCKET NUMBER: 117-181
; TELEPHONE: (703) 816-4000
; TELEFAX: (703) 816-4100
; INFORMATION FOR SEQ ID NO: 13:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 17 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; MISMATCH TYPE: protein
US-09-078-596-13
Query Match      100.0%; Score 31; DB 3; Length 17;
Best Local Similarity 100.0%; Pred. No. 0.62; Indels 0; Gaps 0;
Matches 7; Conservative 0; Mismatches 0;

Qy 1 ALNVLMA 7
Db 6 ALNVLMA 12

RESULT 5
US-09-078-596-13
; Sequence 13, Application US/09078596
; Patent No. 6150116
; GENERAL INFORMATION:
; APPLICANT: La Thangue, Nicholas Barrie
; TITLE OF INVENTION: Transcription Factor DP-1
; NUMBER OF SEQUENCES: 14
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Nixon & Vanderhye
; STREET: 1100 No. 6150116th Glebe Road, 8th Floor
; CITY: Arlington
; STATE: Virginia
; COUNTRY: U.S.A.
; ZIP: 22201-4714
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: IBM PC compatible
; SOFTWARE: PatentIn Release #1.0, Version #1.25 (EPO)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/078,596
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/428,131
; FILING DATE: 23-JUN-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Arthur R. Crawford
; REGISTRATION NUMBER: 25,327
; REFERENCE/DOCKET NUMBER: 117-181
; TELEPHONE: (703) 816-4000
; TELEFAX: (703) 816-4100
; INFORMATION FOR SEQ ID NO: 13:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 17 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; MISMATCH TYPE: protein
US-09-078-596-13
Query Match      100.0%; Score 31; DB 3; Length 17;
Best Local Similarity 100.0%; Pred. No. 0.62; Indels 0; Gaps 0;
Matches 7; Conservative 0; Mismatches 0;

Qy 1 ALNVLMA 7
Db 6 ALNVLMA 12

RESULT 5
US-09-078-596-13
; Sequence 13, Application US/09078596
; Patent No. 6150116
; GENERAL INFORMATION:
; APPLICANT: La Thangue, Nicholas Barrie
; TITLE OF INVENTION: Transcription Factor DP-1
; NUMBER OF SEQUENCES: 14
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Nixon & Vanderhye
; STREET: 1100 No. 6150116th Glebe Road, 8th Floor
; CITY: Arlington
; STATE: Virginia
; COUNTRY: U.S.A.
; ZIP: 22201-4714
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: IBM PC compatible
; SOFTWARE: PatentIn Release #1.0, Version #1.25 (EPO)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/078,596
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/428,131
; FILING DATE: 23-JUN-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Arthur R. Crawford
; REGISTRATION NUMBER: 25,327
; REFERENCE/DOCKET NUMBER: 117-181
; TELEPHONE: (703) 816-4000
; TELEFAX: (703) 816-4100
; INFORMATION FOR SEQ ID NO: 13:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 17 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; MISMATCH TYPE: protein
US-09-078-596-13
Query Match      100.0%; Score 31; DB 3; Length 17;
Best Local Similarity 100.0%; Pred. No. 0.62; Indels 0; Gaps 0;
Matches 7; Conservative 0; Mismatches 0;

Qy 1 ALNVLMA 7
Db 6 ALNVLMA 12

RESULT 5
US-09-078-596-13
; Sequence 13, Application US/09078596
; Patent No. 6150116
; GENERAL INFORMATION:
; APPLICANT: La Thangue, Nicholas Barrie
; TITLE OF INVENTION: Transcription Factor DP-1
; NUMBER OF SEQUENCES: 14
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Nixon & Vanderhye
; STREET: 1100 No. 6150116th Glebe Road, 8th Floor
; CITY: Arlington
; STATE: Virginia
; COUNTRY: U.S.A.
; ZIP: 22201-471
```

Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ALNVLM 7  
| | | | |  
Db 6 ALNVLM 12

## RESULT 6

US-09-308-935-3  
; Sequence 3, Application US/09308935  
; Patent No. 6268334  
; GENERAL INFORMATION:  
; APPLICANT: La Thangue, Nicholas B  
; APPLICANT: Bandara, Lasantha R  
; TITLE OF INVENTION: Peptide antagonists of DP transcription factors  
; FILE REFERENCE: 620-67  
; CURRENT APPLICATION NUMBER: US/09/308,935  
; EARLIER FILING DATE: 1999-05-27  
; EARLIER APPLICATION NUMBER: PCT/GB97/03506  
; EARLIER FILING DATE: 1997-12-22  
; EARLIER APPLICATION NUMBER: GB 9626589.7  
; NUMBER OF SEQ ID NOS: 18  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 3  
; LENGTH: 19  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Description of Artificial Sequence: Synthetic  
US-09-308-935-3

Query Match 100.0%; Score 31; DB 3; Length 19;  
Best Local Similarity 100.0%; Pred. No. 0.7;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ALNVLM 7  
| | | | |  
Db 7 ALNVLM 13

## RESULT 7

US-09-308-935-16  
; Sequence 16, Application US/09308935  
; Patent No. 6268334  
; GENERAL INFORMATION:  
; APPLICANT: La Thangue, Nicholas B  
; APPLICANT: Bandara, Lasantha R  
; TITLE OF INVENTION: Peptide antagonists of DP transcription factors  
; FILE REFERENCE: 620-67  
; CURRENT APPLICATION NUMBER: US/09/308,935  
; EARLIER FILING DATE: 1999-05-27  
; EARLIER APPLICATION NUMBER: PCT/GB97/03506  
; EARLIER FILING DATE: 1997-12-22  
; EARLIER APPLICATION NUMBER: GB 9626589.7  
; NUMBER OF SEQ ID NOS: 18  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 16  
; LENGTH: 19  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Description of Artificial Sequence: Mutant peptide  
US-09-308-935-16

Query Match 100.0%; Score 31; DB 3; Length 19;  
Best Local Similarity 100.0%; Pred. No. 0.7;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ALNVLM 7  
| | | | |  
Db 7 ALNVLM 13

## RESULT 8

US-09-308-935-6  
; Sequence 6, Application US/09308935  
; Patent No. 6268334  
; GENERAL INFORMATION:  
; APPLICANT: La Thangue, Nicholas B  
; APPLICANT: Bandara, Lasantha R  
; TITLE OF INVENTION: Peptide antagonists of DP transcription factors  
; FILE REFERENCE: 620-67  
; CURRENT APPLICATION NUMBER: US/09/308,935  
; EARLIER FILING DATE: 1999-05-27  
; EARLIER APPLICATION NUMBER: PCT/GB97/03506  
; EARLIER FILING DATE: 1997-12-22  
; EARLIER APPLICATION NUMBER: GB 9626589.7  
; NUMBER OF SEQ ID NOS: 18  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 6  
; LENGTH: 30  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Description of Artificial Sequence: Synthetic  
US-09-308-935-6

Query Match 100.0%; Score 31; DB 3; Length 30;  
Best Local Similarity 100.0%; Pred. No. 1.2;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ALNVLM 7  
| | | | |  
Db 3 ALNVLM 9

## RESULT 9

US-09-308-935-1  
; Sequence 1, Application US/09308935  
; Patent No. 6268334  
; GENERAL INFORMATION:  
; APPLICANT: La Thangue, Nicholas B  
; APPLICANT: Bandara, Lasantha R  
; TITLE OF INVENTION: Peptide antagonists of DP transcription factors  
; FILE REFERENCE: 620-67  
; CURRENT APPLICATION NUMBER: US/09/308,935  
; EARLIER FILING DATE: 1999-05-27  
; EARLIER APPLICATION NUMBER: PCT/GB97/03506  
; EARLIER FILING DATE: 1997-12-22  
; EARLIER APPLICATION NUMBER: GB 9626589.7  
; NUMBER OF SEQ ID NOS: 18  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 1  
; LENGTH: 37  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Description of Artificial Sequence: Synthetic  
US-09-308-935-1

Query Match 100.0%; Score 31; DB 3; Length 37;  
Best Local Similarity 100.0%; Pred. No. 1.5;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ALNVLM 7  
| | | | |  
Db 10 ALNVLM 16

## RESULT 10

US-08-428-131-11  
; Sequence 11, Application US/08428131

us-09-900-147-7-rai

Tue Feb 17 11:55:56 2004

```

; Patent No. 5863757
; GENERAL INFORMATION:
; APPLICANT: La Thangue, Nicholas Barrie
; TITLE OF INVENTION: Transcription Factor DP-1
; NUMBER OF SEQUENCES: 14
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Nixon & Vanderhye
; STREET: 1100 No. 5863757th Glebe Road, 8th Floor
; CITY: Arlington
; STATE: Virginia
; COUNTRY: U.S.A.
; ZIP: 22201-4714
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25 (EPO)
; CURRENT APPLICATION NUMBER: US/08/428,131
; FILING DATE: 23-JUN-1995
; CLASSIFICATION: 514
; ATTORNEY/AGENT INFORMATION:
; NAME: Arthur R. Crawford
; REGISTRATION NUMBER: 25,327
; REFERENCE/DOCKET NUMBER: 117-181
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703) 816-4000
; TELEFAX: (703) 816-4100
; INFORMATION FOR SEQ ID NO: 11:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 72 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-09-078-596-11

Query Match 100.0%; Score 31; DB 3; Length 72;
Best Local Similarity 100.0%; Pred. No. 3;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ALNVLMA 7
Db 13 ALNVLMA 19

RESULT 12
US-08-894-139-10
; Sequence 10, Application US/08894139
; Patent No. 6448376
; GENERAL INFORMATION:
; APPLICANT: LA THANGUE, NICHOLAS B.
; APPLICANT: BERNARDS, RENE
; APPLICANT: HIGMANS, ELEANORE M.
; TITLE OF INVENTION: TRANSCRIPTION FACTOR E2P-5
; NUMBER OF SEQUENCES: 25
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: NIXON & VANDERHYE P.C.
; STREET: 1100 NORTH GLEBE ROAD
; CITY: ARLINGTON
; STATE: VIRGINIA
; COUNTRY: U.S.A.
; ZIP: 22201-4714
; COMPUTER READABLE FORM: disk
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/894,139
; FILING DATE: 13-AUG-1997
; CLASSIFICATION: 536
; ATTORNEY/AGENT INFORMATION:
; NAME: WILSON, MARY J.
; REGISTRATION NUMBER: 32,955
; REFERENCE/DOCKET NUMBER: 620-22
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703) 816-4000
; TELEFAX: (703) 816-4100
; INFORMATION FOR SEQ ID NO: 10:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 74 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; US-08-894-139-10

Query Match 100.0%; Score 31; DB 4; Length 74;
Best Local Similarity 100.0%; Pred. No. 3.1;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ALNVLMA 7
Db 13 ALNVLMA 19

; Patent No. 5863757
; GENERAL INFORMATION:
; APPLICANT: La Thangue, Nicholas Barrie
; TITLE OF INVENTION: Transcription Factor DP-1
; NUMBER OF SEQUENCES: 14
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Nixon & Vanderhye
; STREET: 1100 No. 6150116th Glebe Road, 8th Floor
; CITY: Arlington
; STATE: Virginia
; COUNTRY: U.S.A.
; ZIP: 22201-4714
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25 (EPO)
; CURRENT APPLICATION NUMBER: US/09/078,596
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/428,131
; FILING DATE: 23-JUN-1995

```

Db 52 ALNVLMA 58

RESULT 13  
US-09-308-935-15  
; Sequence 15, Application US/09308935  
; Patent No. 6268334  
; GENERAL INFORMATION:  
; APPLICANT: La Thangue, Nicholas B  
; TITLE OF INVENTION: Peptide antagonists of DP transcription factors  
; FILE REFERENCE: 620-67  
; CURRENT APPLICATION NUMBER: US/09/308,935  
; EARLIER FILING DATE: 1999-05-27  
; EARLIER APPLICATION NUMBER: PCT/GB97/03506  
; EARLIER FILING DATE: 1997-12-22  
; EARLIER APPLICATION NUMBER: GB 9626589.7  
; EARLIER FILING DATE: 1996-12-20  
; NUMBER OF SEQ ID NOS: 18  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 15  
; LENGTH: 19  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Description of Artificial Sequence: Mutant peptide  
US-09-308-935-15

Query Match 87.1%; Score 27; DB 3; Length 19;  
Best Local Similarity 85.7%; Pred. No. 5.7;  
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 ALNVLMA 7  
||| |||  
Db 7 ALNALMA 13

RESULT 14  
US-09-152-060-89  
; Sequence 89, Application US/09152060  
; Patent No. 6448230  
; GENERAL INFORMATION:  
; APPLICANT: Rosen et al.  
; TITLE OF INVENTION: 28 Human Secreted Proteins  
; FILE REFERENCE: P2003P1.US  
; CURRENT APPLICATION NUMBER: US/09/152,060  
; CURRENT FILING DATE: 1998-09-11  
; EARLIER APPLICATION NUMBER: PCT/US98/04858  
; EARLIER FILING DATE: 1998-03-12  
; EARLIER APPLICATION NUMBER: 60/040,762  
; EARLIER FILING DATE: 1997-03-14  
; EARLIER APPLICATION NUMBER: 60/040,710  
; EARLIER FILING DATE: 1997-03-14  
; EARLIER APPLICATION NUMBER: 60/050,934  
; EARLIER FILING DATE: 1997-05-30  
; EARLIER APPLICATION NUMBER: 60/048,100  
; EARLIER FILING DATE: 1997-05-30  
; EARLIER APPLICATION NUMBER: 60/048,357  
; EARLIER FILING DATE: 1997-05-30  
; EARLIER APPLICATION NUMBER: 60/048,189  
; EARLIER FILING DATE: 1997-05-30  
; EARLIER APPLICATION NUMBER: 60/057,765  
; EARLIER FILING DATE: 1997-09-05  
; EARLIER APPLICATION NUMBER: 60/048,970  
; EARLIER FILING DATE: 1997-06-06  
; EARLIER APPLICATION NUMBER: 60/068,368  
; EARLIER FILING DATE: 1997-12-19  
; NUMBER OF SEQ ID NOS: 118  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 89  
; LENGTH: 87  
; TYPE: PRT  
; ORGANISM: Homo sapiens

Query Match 80.6%; Score 25; DB 4; Length 87;  
Best Local Similarity 71.4%; Pred. No. 84;  
Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 ALNVLMA 7  
:||||:  
Db 2 SLNVLLA 8

FEATURE:  
; NAME/KEY: SITE  
; LOCATION: (11)  
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids  
; FEATURE:  
; NAME/KEY: SITE  
; LOCATION: (86)  
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids  
US-09-152-060-89

Query Match 80.6%; Score 25; DB 4; Length 87;  
Best Local Similarity 71.4%; Pred. No. 84;  
Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 ALNVLMA 7  
:||||:  
Db 2 SLNVLLA 8

RESULT 15  
US-08-484-493-5  
; Sequence 5, Application US/08484493  
; Patent No. 5728381  
; GENERAL INFORMATION:  
; APPLICANT: Wilson, Peter J  
; APPLICANT: Morris, Charles P  
; APPLICANT: Anson, Donald S  
; APPLICANT: Ochiodoro, Teresa  
; APPLICANT: Bielicki, Julie  
; APPLICANT: Clements, Peter R  
; APPLICANT: Hopwood, John J  
; TITLE OF INVENTION: GLYCOSYLATION VARIANTS OF  
; TITLE OF INVENTION: IDURONATE 2-SULFATASE  
; NUMBER OF SEQUENCES: 15  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Scully, Scott, Murphy & Presser  
; STREET: 400 Garden City Plaza  
; CITY: Garden City  
; STATE: New York  
; COUNTRY: USA  
; ZIP: 11530  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/484,493  
; FILING DATE: 07-JUN-1995  
; CLASSIFICATION: 424  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 991,973  
; FILING DATE: 17-DEC-1992  
; ATTORNEY/AGENT INFORMATION:  
; NAME: DiGiglio, Frank S  
; REGISTRATION NUMBER: 31,346  
; REFERENCE/DOCKET NUMBER: 8416Z  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 516-742-4343  
; TELEFAX: 516-742-4366  
; TELEX: 230 901 SANS UR  
; INFORMATION FOR SEQ ID NO: 5:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 25 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: peptide  
US-08-484-493-5

Query Match 77.4%; Score 24; DB 1; Length 25;  
Best Local Similarity 83.3%; Pred. No. 36;  
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 ALNVLM 6  
Db 3 ALNVLL 8

RESULT 16  
US-08-484-494-5  
; Sequence 5, Application US/08484494  
; Patent No. 5798239  
; GENERAL INFORMATION:  
; APPLICANT: Wilson, Peter J  
; APPLICANT: Morris, Charles P  
; APPLICANT: Anson, Donald S  
; APPLICANT: Occhiodoro, Teresa  
; APPLICANT: Bielicki, Julie  
; APPLICANT: Clements, Peter R  
; APPLICANT: Hopwood, John J  
; TITLE OF INVENTION: GLYCOSYLATION VARIANTS OF  
; IDURONATE 2-SULFATASE  
; NUMBER OF SEQUENCES: 15  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Scully, Scott, Murphy & Presser  
; STREET: 400 Garden City Plaza  
; CITY: Garden City  
; STATE: New York  
; COUNTRY: USA  
; ZIP: 11530  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent In Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/345,212  
; FILING DATE:  
; CLASSIFICATION: 435  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 991,973  
; FILING DATE: 17-DEC-1992  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Digiglio, Frank S  
; REGISTRATION NUMBER: 31,346  
; REFERENCE/DOCKET NUMBER: 84162  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 516-742-4343  
; TELEFAX: 516-742-4366  
; TELEX: 230 901 SANS UR  
; INFORMATION FOR SEQ ID NO: 5:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 25 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: peptide  
; US-08-345-212-5

Query Match 77.4%; Score 24; DB 2; Length 25;  
Best Local Similarity 83.3%; Pred. No. 36;  
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 ALNVLM 6  
Db 3 ALNVLL 8

RESULT 18  
US-09-249-003-5  
; Sequence 5, Application US/09249003  
; Patent No. 6153188  
; GENERAL INFORMATION:  
; APPLICANT: Wilson, Peter J  
; APPLICANT: Morris, Charles P  
; APPLICANT: Anson, Donald S  
; APPLICANT: Occhiodoro, Teresa  
; APPLICANT: Bielicki, Julie  
; APPLICANT: Clements, Peter R  
; APPLICANT: Hopwood, John J  
; TITLE OF INVENTION: GLYCOSYLATION VARIANTS OF  
; IDURONATE 2-SULFATASE  
; NUMBER OF SEQUENCES: 15  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Scully, Scott, Murphy & Presser  
; STREET: 400 Garden City Plaza  
; CITY: Garden City  
; STATE: New York

QY 1 ALNVLM 6  
Db 3 ALNVLL 8

RESULT 17  
US-08-345-212-5  
; Sequence 5, Application US/08345212  
; Patent No. 5932211  
; GENERAL INFORMATION:  
; APPLICANT: Wilson, Peter J  
; APPLICANT: Morris, Charles P

```
;
; COUNTRY: USA
; ZIP: 11530
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/249,003
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 07/991,973
; FILING DATE: 17-DEC-1992
; ATTORNEY/AGENT INFORMATION:
; NAME: Digiglio, Frank S
; REGISTRATION NUMBER: 31,346
; REFERENCE/DOCKET NUMBER: 8416Z
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 516-742-4343
; TELEFAX: 516-742-4366
; TELEX: 230 901 SANS UR
; INFORMATION FOR SEQ ID NO: 5:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 25 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; US-09-249-003-5

Query Match 77.4%; Score 24; DB 3; Length 25;
Best Local Similarity 83.3%; Pred. No. 36;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ALNVLM 6
Db 3 ALNVLL 8

RESULT 19
US-09-685-844-5
; Sequence 5, Application US/09685844
; Patent No. 6541254
; GENERAL INFORMATION:
; APPLICANT: Wilson, Peter J
; Morris, Charles P
; Anson, Donald S
; Occhiodoro, Teresa
; Bielicki, Julie
; Clements, Peter R
; Hopwood, John J
; TITLE OF INVENTION: GLYCOSYLATION VARIANTS OF
; IDURONATE 2-SULFATASE
; NUMBER OF SEQUENCES: 15
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Scully, Scott, Murphy & Presser
; STREET: 400 Garden City Plaza
; CITY: Garden City
; STATE: New York
; COUNTRY: USA
; ZIP: 11530
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/685,844
; FILING DATE: 10-Oct-2000
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 07/991,973
```

```
;
; FILING DATE: 17-DEC-1992
; ATTORNEY/AGENT INFORMATION:
; NAME: Digiglio, Frank S
; REGISTRATION NUMBER: 31,346
; REFERENCE/DOCKET NUMBER: 8416Z
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 516-742-4343
; TELEFAX: 516-742-4366
; TELEX: 230 901 SANS UR
; INFORMATION FOR SEQ ID NO: 5:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 25 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; SEQUENCE DESCRIPTION: SEQ ID NO: 5:
US-09-685-844-5

Query Match 77.4%; Score 24; DB 4; Length 25;
Best Local Similarity 83.3%; Pred. No. 36;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ALNVLM 6
Db 3 ALNVLL 8

RESULT 20
US-08-194-338-14
; Sequence 14, Application US/08194338
; Patent No. 5474898
; GENERAL INFORMATION:
; APPLICANT: Venter, John C.
; APPLICANT: Fraser, Claire M.
; APPLICANT: McCombie, William R.
; TITLE OF INVENTION: OCTOPAMINE RECEPTOR
; NUMBER OF SEQUENCES: 16
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Knobbe, Martens, Olson and Bear
; STREET: 620 Newport Center Drive, Sixteenth Floor
; CITY: Newport Beach
; STATE: CA
; COUNTRY: USA
; ZIP: 92660
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/194,338
; FILING DATE: 08-FEB-1994
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/676,174
; FILING DATE: 28-MAR-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: Israelsson, Ned A.
; REGISTRATION NUMBER: 29,655
; REFERENCE/DOCKET NUMBER: NIH101.001DV1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (619) 235-8550
; TELEFAX: (619) 235-0176
; INFORMATION FOR SEQ ID NO: 14:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 63 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
```

```

RESULT 22
US-09-308-935-4
; Sequence 4, Application US/09308935
; Patent No. 6268334
; GENERAL INFORMATION:
; APPLICANT: La Thangue, Nicholas B
; APPLICANT: Bandara, Lasantha R
; TITLE OF INVENTION: Peptide antagonists of DP transcription factors
; FILE REFERENCE: 620-67
; CURRENT APPLICATION NUMBER: US/09/308,935
; CURRENT FILING DATE: 1999-05-27
; EARLIER APPLICATION NUMBER: PCT/GB97/03506
; EARLIER FILING DATE: 1997-12-22
; EARLIER APPLICATION NUMBER: GB 9626589.7
; EARLIER FILING DATE: 1996-12-20
; NUMBER OF SEQ ID NOS: 18
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 4
; LENGTH: 20
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURES:
; OTHER INFORMATION: Description of Artificial Sequence: Synthetic
; US-09-308-933-4

Query Match          74.2%; Score 23; DB 3; Length 20;
Best Local Similarity 100.0%; Pred. No. 48;

```

RESULT 24  
US-08-811-492-33  
Sequence 33, Application US/08011492  
Patent No. 5834247  
GENERAL INFORMATION:  
APPLICANT: COMB, DONALD G.  
APPLICANT: PERLER, FRANCINE B.  
APPLICANT: JACK, WILLIAM E.  
APPLICANT: XU, MING-QUN  
APPLICANT: HODGES, ROBERT A.  
APPLICANT: NORDEN, CHRISTOPHER J.



APPLICANT: CHONG, SHAORONG S.C.  
APPLICANT: ADAM, ERIC  
APPLICANT: SOUTHWORTH, MAURICE  
TITLE OF INVENTION: MODIFIED PROTEINS, METHODS OF THEIR  
PRODUCTION AND METHODS FOR PURIFICATION OF TARGET  
NUMBER OF SEQUENCES: 155  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: GREGORY D. WILLIAMS; NEW ENGLAND BIOLABS, INC.  
STREET: 32 TOZER ROAD  
CITY: BEVERLY  
STATE: MASSACHUSETTS  
COUNTRY: USA  
ZIP: 01915  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC\ DOS\MS\ DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION NUMBER: US/08/811.492  
FILING DATE:  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/580,555  
FILING DATE: 29-DEC-1995  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/496,247  
FILING DATE: 28-JUN-1995  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/146,885  
FILING DATE: 03-NOV-1993  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/004,139  
FILING DATE: 09-DEC-1992  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: Williams, Gregory D  
REGISTRATION NUMBER: 30901  
REFERENCE/DOCKET NUMBER: NEB-036C4  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 508-927-5054  
TELEFAX: 509-927-1705  
TELEX:  
INFORMATION FOR SEQ ID NO: 33:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 22 amino acids  
TYPE: amino acid  
STRANDEDNESS: unknown  
TOPOLOGY: unknown  
US-08-811-492-33  
Query Match 74.2%; Score 23; DB 2; Length 22;  
Best Local Similarity 100.0%; Pred. No. 53;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Qy 3 NVLMA 7  
Db 13 NVLMA 17  
RESULT 25  
PCT-US96-10545A-33  
GENERAL INFORMATION:  
APPLICANT: COMB, DONALD G.  
APPLICANT: PERLER, FRANCINE B.  
APPLICANT: JACK, WILLIAM E.  
APPLICANT: XU, MING-QUN  
APPLICANT: HODGES, ROBERT A.

APPLICANT: NOREN, CHRISTOPHER J.  
TITLE OF INVENTION: MODIFIED PROTEINS AND METHODS OF THEIR  
PRODUCTION  
NUMBER OF SEQUENCES: 77  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: GREGORY D. WILLIAMS; NEW ENGLAND BIOLABS, INC.  
STREET: 32 TOZER ROAD  
CITY: BEVERLY  
STATE: MASSACHUSETTS  
COUNTRY: USA  
ZIP: 01915  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION NUMBER: PCT/US96/10545A  
FILING DATE:  
CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/580,555  
FILING DATE: 29-DEC-1995  
CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/496,247  
FILING DATE: 28-JUN-1995  
CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/146,885  
FILING DATE: 03-NOV-1993  
CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/004,139  
FILING DATE: 09-DEC-1992  
CLASSIFICATION:  
ATTORNEY/AGENT INFORMATION:  
NAME: WILLIAMS, GREGORY D.  
REGISTRATION NUMBER: 30901  
REFERENCE/DOCKET NUMBER: NEB-036C2  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (508) 927-5054  
TELEFAX: (508) 927-1705  
TELEX:  
INFORMATION FOR SEQ ID NO: 33:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 22 amino acids  
TYPE: amino acid  
STRANDEDNESS: unknown  
TOPOLOGY: unknown  
PCT-US96-10545A-33  
Query Match 74.2%; Score 23; DB 5; Length 22;  
Best Local Similarity 100.0%; Pred. No. 53;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Qy 3 NVLMA 7  
Db 13 NVLMA 17  
RESULT 26  
US-09-215-221-18  
Sequence 18, Application US/09215221  
Patent No. 6265562  
GENERAL INFORMATION:  
APPLICANT: EILERS, MARTIN  
APPLICANT: BUERGIN, ANDREA  
APPLICANT: SEDLACEK, HANS-HARALD  
TITLE OF INVENTION: NUCLEIC ACID CONSTRUCTS WHOSE ACTIVITY IS AFFECTED BY  
INHIBITORS OF CYCLIN-DEPENDANT KINASES AND USES THEREOF  
FILE REFERENCE: 026083/0192  
CURRENT APPLICATION NUMBER: US/09/215,221

OTHER INFORMATION: Description of Artificial Sequence: Synthetic  
US-09-308-935-9

Query Match 71.0%; Score 22; DB 3; Length 11;  
Best Local Similarity 100.0%; Pred. No. 42; 0; Indels 0; Gaps 0;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ALNVL 5  
DB 7 ALNVL 11

RESULT 29  
US-08-725-459B-11  
Sequence 11, Application US/08725459B  
Patent No. 6084068  
GENERAL INFORMATION:  
APPLICANT: CONAWAY, RONALD C.  
APPLICANT: CONAWAY, JOAN W.  
TITLE OF INVENTION: ELONGIN A AND C FUNCTIONAL DOMAINS  
NUMBER OF SEQUENCES: 79  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: SIDLEY & AUSTIN  
STREET: 717 N. HARWOOD, SUITE 3400  
CITY: DALLAS  
STATE: TX  
COUNTRY: US  
ZIP: 75201-6507  
COMPUTER READABLE FORM:  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/725,459B  
FILING DATE: 04-OCT-1996  
CLASSIFICATION: 530  
ATTORNEY/AGENT INFORMATION:  
NAME: HANSEN, EUGENIA S.  
REGISTRATION NUMBER: 31,966  
REFERENCE/DOCKET NUMBER: 11146/07501  
TELEPHONE: 214-981-3300  
TELEFAX: 214-981-3400  
INFORMATION FOR SEQ ID NO: 11:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 13 amino acids  
TYPE: amino acid  
STRANDEDNESS: not relevant  
TOPOLOGY: not relevant  
MOLECULE TYPE: peptide  
FEATURE:  
NAME/KEY: Protein  
LOCATION: 1..13  
OTHER INFORMATION: /note= "amino acids 100-112 of  
OTHER INFORMATION: Elongin C"

Query Match 71.0%; Score 22; DB 3; Length 13;  
Best Local Similarity 71.4%; Pred. No. 51;  
Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 ALNVLMA 7  
DB 1 ALELLMA 7

RESULT 30  
US-09-399-494-19  
Sequence 19, Application US/09399494  
Patent No. 6228563  
GENERAL INFORMATION:  
APPLICANT: Naparstek, Yaakov

CURRENT FILING DATE: 1998-12-18  
PRIOR APPLICATION NUMBER: 197 56 975.7  
PRIOR FILING DATE: 1997-12-20  
NUMBER OF SEQ ID NOS: 57  
SOFTWARE: PatentIn Ver. 2.1  
SEQ ID NO 18  
LENGTH: 40  
TYPE: PRT  
ORGANISM: Murine sp.  
US-09-215-221-18

Query Match 74.2%; Score 23; DB 3; Length 40;  
Best Local Similarity 83.3%; Pred. No. 1e+02; 0; Indels 0; Gaps 0;  
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 2 LNVLMMA 7  
DB 6 LNVLLIA 11

RESULT 27  
US-09-134-001C-4216  
Sequence 4216, Application US/09134001C  
Patent No. 6380370  
GENERAL INFORMATION:  
APPLICANT: Lynn Doucette-Stamm et al  
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO STAPHYLOCOCCUS  
TITLE OF INVENTION: EPIDERMIDIS FOR DIAGNOSTICS AND THERAPEUTICS  
FILE REFERENCE: GTC-007  
CURRENT APPLICATION NUMBER: US/09/134,001C  
CURRENT FILING DATE: 1998-08-13  
PRIOR APPLICATION NUMBER: US 60/064,964  
PRIOR FILING DATE: 1997-11-08  
PRIOR APPLICATION NUMBER: US 60/055,779  
PRIOR FILING DATE: 1997-08-14  
NUMBER OF SEQ ID NOS: 5674  
SEQ ID NO 4216  
LENGTH: 67  
TYPE: PRT  
ORGANISM: Staphylococcus epidermidis  
US-09-134-001C-4216

Query Match 74.2%; Score 23; DB 4; Length 67;  
Best Local Similarity 66.7%; Pred. No. 1.8e+02;  
Matches 4; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 2 LNVLMMA 7  
DB 15 LNVLLMS 20

RESULT 28  
US-09-308-935-9  
Sequence 9, Application US/09308935  
Patent No. 6268334  
GENERAL INFORMATION:  
APPLICANT: La Thangue, Nicholas B  
APPLICANT: Bandara, Lasantha R  
TITLE OF INVENTION: Peptide antagonists of DP transcription factors  
FILE REFERENCE: 620-67  
CURRENT APPLICATION NUMBER: US/09/308,935  
CURRENT FILING DATE: 1999-05-27  
EARLIER APPLICATION NUMBER: PCT/GB97/03506  
EARLIER FILING DATE: 1997-12-22  
EARLIER APPLICATION NUMBER: GB 9626589.7  
EARLIER FILING DATE: 1996-12-20  
NUMBER OF SEQ ID NOS: 18  
SOFTWARE: PatentIn Ver. 2.1  
SEQ ID NO 9  
LENGTH: 11  
TYPE: PRT  
ORGANISM: Artificial Sequence  
FEATURE:

;; TITLE OF INVENTION: PEPTIDES FOR THE TREATMENT OF SYSTEMIC LUPUS

;; FILE REFERENCE: ERYTHEMATOSUS

;; CURRENT FILING DATE: 1999-09-20

;; NUMBER OF SEQ ID NOS: 22

;; SOFTWARE: Patentin Ver. 2.1

;; SEQ ID NO 19

;; LENGTH: 21

;; TYPE: PRT

;; ORGANISM: mouse

US-09-399-494-19

Query Match 71.0%; Score 22; DB 3; Length 21;

Best Local Similarity 57.1%; Pred. No. 85;

Matches 4; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Qy 1 ALNVLMA 7

Db 9 ALNITLA 15

RESULT 31

US-08-725-459B-13

; Sequence 13, Application US/08725459B

; Patent No. 6084068

; GENERAL INFORMATION:

; APPLICANT: CONAWAY, RONALD C.

; APPLICANT: CONAWAY, JOAN W.

; TITLE OF INVENTION: ELONGIN A AND C FUNCTIONAL DOMAINS

; NUMBER OF SEQUENCES: 79

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: SIDLEY & AUSTIN

; STREET: 717 N. HARWOOD, SUITE 3400

; CITY: DALLAS

; STATE: TX

; COUNTRY: US

; ZIP: 75201-6507

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: Patentin Release #1.0, Version #1.30

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/725,459B

; FILING DATE: 04-OCT-1996

; CLASSIFICATION: 530

; ATTORNEY/AGENT INFORMATION:

; NAME: HANSEN, EUGENIA S.

; REGISTRATION NUMBER: 31,966

; REFERENCE/DOCKET NUMBER: 11146/07501

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: 214-981-3300

; TELEFAX: 214-981-3400

; INFORMATION FOR SEQ ID NO: 13:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 22 amino acids

; TYPE: amino acid

; STRANDEDNESS: not relevant

; TOPOLOGY: not relevant

; MOLECULE TYPE: peptide

; FEATURE:

; NAME/KEY: Peptide

; LOCATION: 1..22

; OTHER INFORMATION: /note= "amino acids 91-112 of

; ELONGIN C"

US-08-725-459B-13

Query Match 71.0%; Score 22; DB 3; Length 22;

Best Local Similarity 71.4%; Pred. No. 90;

Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 1 ALNVLMA 7

Db 10 ALELLMA 16

RESULT 32

US-08-725-459B-75

; Sequence 75, Application US/08725459B

; Patent No. 6084068

; GENERAL INFORMATION:

; APPLICANT: CONAWAY, RONALD C.

; APPLICANT: CONAWAY, JOAN W.

; TITLE OF INVENTION: ELONGIN A AND C FUNCTIONAL DOMAINS

; NUMBER OF SEQUENCES: 79

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: SIDLEY & AUSTIN

; STREET: 717 N. HARWOOD, SUITE 3400

; CITY: DALLAS

; STATE: TX

; COUNTRY: US

; ZIP: 75201-6507

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: Patentin Release #1.0, Version #1.30

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/725,459B

; FILING DATE: 04-OCT-1996

; CLASSIFICATION: 530

; ATTORNEY/AGENT INFORMATION:

; NAME: HANSEN, EUGENIA S.

; REGISTRATION NUMBER: 31,966

; REFERENCE/DOCKET NUMBER: 11146/07501

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: 214-981-3300

; TELEFAX: 214-981-3400

; INFORMATION FOR SEQ ID NO: 75:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 24 amino acids

; TYPE: amino acid

; STRANDEDNESS: not relevant

; TOPOLOGY: not relevant

; MOLECULE TYPE: peptide

; FEATURE:

; NAME/KEY: Peptide

; LOCATION: 1..24

; OTHER INFORMATION: /note= "amino acids 89-112 of

; ELONGIN C"

US-08-725-459B-75

Query Match 71.0%; Score 22; DB 3; Length 24;

Best Local Similarity 71.4%; Pred. No. 99;

Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 1 ALNVLMA 7

Db 12 ALELLMA 18

RESULT 33

US-08-524-757-16

; Sequence 16, Application US/08524757

; Patent No. 5792634

; GENERAL INFORMATION:

; APPLICANT: Conaway, Ronald C.

; APPLICANT: Conaway, Joan W.

; TITLE OF INVENTION: RNA Polymerase Transcription Factor

; NUMBER OF SEQUENCES: 37

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: RICHARDS, MEDLOCK & ANDREWS

; STREET: 1201 Elm Street, Suite 4500

; CITY: Dallas

Qy 1 ALNVLMA 7

Ov 1 ALNVLM 6

STATE: MA  
COUNTRY: USA  
ZIP: 02109-2891  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/152.721B  
FILING DATE: 15-NOV-1993  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: Pasternack Esq., Sam  
REGISTRATION NUMBER: 29,576  
REFERENCE/DOCKET NUMBER: 181411-011DIV  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (617) 248-5000  
TELEFAX: (617) 248-4000  
INFORMATION FOR SEQ ID NO: 22:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 55 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
HYPOTHETICAL: NO  
ANTI-SENSE: NO  
FRAGMENT TYPE: internal  
IMMEDIATE SOURCE:  
CLONE: RB region 3  
US-08-152-721B-22

Query Match 71.0%; Score 22; DB 2; Length 55;  
Best Local Similarity 71.4%; Pred. No. 2.4e+02;  
Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 1 ALNVLMA 7  
|||  
Db 46 ALEVVMA 52

RESULT 37  
US-08-725-459B-50  
Sequence 50, Application US/08725459B  
Patent No. 6084068  
GENERAL INFORMATION:  
APPLICANT: CONAWAY, RONALD C.  
TITLE OF INVENTION: ELONGIN A AND C FUNCTIONAL DOMAINS  
NUMBER OF SEQUENCES: 79  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: SIDLEY & AUSTIN  
STREET: 717 N. HARMWOOD, SUITE 3400  
CITY: DALLAS  
STATE: TX  
COUNTRY: US  
ZIP: 75201-6507  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/725.459B  
FILING DATE: 04-OCT-1996  
CLASSIFICATION: 530  
ATTORNEY/AGENT INFORMATION:  
NAME: HANSEN, EUGENIA S.  
REGISTRATION NUMBER: 31,966  
REFERENCE/DOCKET NUMBER: 11146/07501  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 214-981-3300

TELEFAX: 214-981-3400  
INFORMATION FOR SEQ ID NO: 50:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 56 amino acids  
TYPE: amino acid  
STRANDEDNESS: not relevant  
TOPOLOGY: not relevant  
MOLECULE TYPE: peptide  
FEATURE:  
NAME/KEY: Peptide  
LOCATION: 1..56  
OTHER INFORMATION: /note= "amino acids 57-112 of Elongin C"  
US-08-725-459B-50

Query Match 71.0%; Score 22; DB 3; Length 56;  
Best Local Similarity 71.4%; Pred. No. 2.5e+02;  
Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 1 ALNVLMA 7  
|||  
Db 44 ALELLMA 50

RESULT 38  
US-09-328-352-5729  
Sequence 5729, Application US/09328352  
Patent No. 6562958  
GENERAL INFORMATION:  
APPLICANT: Gary L. Breton et al.  
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO ACINETOBACTER  
FILE REFERENCE: GTC99-03PA  
CURRENT APPLICATION NUMBER: US/09/328.352  
CURRENT FILING DATE: 1999-06-04  
NUMBER OF SEQ ID NOS: 8252  
SEQ ID NO 5729  
LENGTH: 63  
TYPE: PRT  
ORGANISM: Acinetobacter baumannii  
US-09-328-352-5729

Query Match 71.0%; Score 22; DB 4; Length 63;  
Best Local Similarity 66.7%; Pred. No. 2.8e+02;  
Matches 4; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Qy 2 LNVLMA 7  
|||  
Db 46 LNLIMA 51

RESULT 39  
US-08-530-010-18  
Sequence 18, Application US/08530010  
Patent No. 5689055  
GENERAL INFORMATION:  
APPLICANT: Meyerowitz, Elliott M.  
APPLICANT: Chang, Caren  
APPLICANT: Blecker, Anthony B.  
TITLE OF INVENTION: PLANTS HAVING MODIFIED RESPONSE TO ETHYLENE  
NUMBER OF SEQUENCES: 34  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Richard F. Trecartin  
STREET: 3400 Embarcadero Center, Suite 3400  
CITY: San Francisco  
STATE: California  
COUNTRY: USA  
ZIP: 94111  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25



REGISTRATION NUMBER: 31,966  
REFERENCE/DOCKET NUMBER: 11146/07501  
TELEPHONE: 214-981-3300  
TELEFAX: 214-981-3400  
INFORMATION FOR SEQ ID NO: 49:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 84 amino acids  
TYPE: amino acid  
STRANDEDNESS: not relevant  
TOPOLOGY: not relevant  
MOLECULE TYPE: peptide  
FEATURE:  
NAME/KEY: Peptide  
LOCATION: 1..84  
OTHER INFORMATION: /note= "amino acids 29-112 of Elongin C"  
US-08-725-459B-49

Query Match 71.0%; Score 22; DB 3; Length 84;  
Best Local Similarity 71.4%; Pred. No. 3.8e+02;  
Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 1 ALNVLMA 7  
Db 72 ALELLMA 78

RESULT 43  
US-08-725-459B-48  
Sequence 48, Application US/08725459B  
Patent No. 6084068  
GENERAL INFORMATION:  
APPLICANT: CONAWAY, RONALD C.  
APPLICANT: CONAWAY, JOAN W.  
TITLE OF INVENTION: ELONGIN A AND C FUNCTIONAL DOMAINS  
NUMBER OF SEQUENCES: 79  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: SIDLEY & AUSTIN  
STREET: 717 N. HARWOOD, SUITE 3400  
CITY: DALLAS  
STATE: TX  
COUNTRY: US  
ZIP: 75201-6507  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/725,459B  
FILING DATE: 04-OCT-1996  
CLASSIFICATION: 530  
ATTORNEY/AGENT INFORMATION:  
NAME: HANSEN, EUGENIA S.  
REGISTRATION NUMBER: 31,966  
REFERENCE/DOCKET NUMBER: 11146/07501  
TELEPHONE: 214-981-3300  
TELEFAX: 214-981-3400  
INFORMATION FOR SEQ ID NO: 48:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 90 amino acids  
TYPE: amino acid  
STRANDEDNESS: not relevant  
TOPOLOGY: not relevant  
MOLECULE TYPE: peptide  
FEATURE:  
NAME/KEY: Peptide  
LOCATION: 1..90  
OTHER INFORMATION: /note= "amino acids 23-112 of Elongin C"  
US-08-725-459B-48

Query Match 71.0%; Score 22; DB 3; Length 90;  
Best Local Similarity 71.4%; Pred. No. 4.1e+02;  
Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 1 ALNVLMA 7  
Db 78 ALELLMA 84

RESULT 44  
US-08-725-459B-47  
Sequence 47, Application US/08725459B  
Patent No. 6084068  
GENERAL INFORMATION:  
APPLICANT: CONAWAY, RONALD C.  
APPLICANT: CONAWAY, JOAN W.  
TITLE OF INVENTION: ELONGIN A AND C FUNCTIONAL DOMAINS  
NUMBER OF SEQUENCES: 79  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: SIDLEY & AUSTIN  
STREET: 717 N. HARWOOD, SUITE 3400  
CITY: DALLAS  
STATE: TX  
COUNTRY: US  
ZIP: 75201-6507  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/725,459B  
FILING DATE: 04-OCT-1996  
CLASSIFICATION: 530  
ATTORNEY/AGENT INFORMATION:  
NAME: HANSEN, EUGENIA S.  
REGISTRATION NUMBER: 31,966  
REFERENCE/DOCKET NUMBER: 11146/07501  
TELEPHONE: 214-981-3300  
TELEFAX: 214-981-3400  
INFORMATION FOR SEQ ID NO: 47:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 94 amino acids  
TYPE: amino acid  
STRANDEDNESS: not relevant  
TOPOLOGY: not relevant  
MOLECULE TYPE: peptide  
FEATURE:  
NAME/KEY: Peptide  
LOCATION: 1..94  
OTHER INFORMATION: /note= "amino acids 19-112 of Elongin C"  
US-08-725-459B-47

Query Match 71.0%; Score 22; DB 3; Length 94;  
Best Local Similarity 71.4%; Pred. No. 4.4e+02;  
Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 1 ALNVLMA 7  
Db 82 ALELLMA 88

RESULT 45  
US-08-725-459B-46  
Sequence 46, Application US/08725459B  
Patent No. 6084068  
GENERAL INFORMATION:  
APPLICANT: CONAWAY, RONALD C.  
APPLICANT: CONAWAY, JOAN W.  
TITLE OF INVENTION: ELONGIN A AND C FUNCTIONAL DOMAINS

Tue Feb 17 11:55:56 2004

```

; NUMBER OF SEQUENCES: 79
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: SIDLEY & AUSTIN
; STREET: 717 N. HARWOOD, SUITE 3400
; CITY: DALLAS
; STATE: TX
; COUNTRY: US
; ZIP: 75201-6507
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/725,459B
; FILING DATE: 04-OCT-1996
; CLASSIFICATION: 530
; ATTORNEY/AGENT INFORMATION:
; NAME: HANSEN, EUGENIA S.
; REGISTRATION NUMBER: 31,966
; REFERENCE/DOCKET NUMBER: 11146/07501
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 214-981-3300
; TELEFAX: 214-981-3400
; INFORMATION FOR SEQ ID NO: 46:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 98 amino acids
; TYPE: amino acid
; STRANDEDNESS: not relevant
; TOPOLOGY: not relevant
; MOLECULE TYPE: peptide
; FEATURE:
; NAME/KEY: Peptide
; LOCATION: 1..98
; OTHER INFORMATION: /note= "amino acids 15-112 of
; US-08-725-459B-46

```

```

Query Match          71.0%; Score 22; DB 3; Length 98;
Best Local Similarity 71.4%; Pred. NO. 4.6e+02;
Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

```

```

Qy      1 ALNVLMA 7
        |||||
Db      86 ALELLMA 92

```

Search completed: February 17, 2004, 10:59:42  
Job time : 6.05941 secs



GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

# OM protein - protein search, using sw model

Run on: February 17, 2004, 10:50:12 ; Search time 9.91089 Seconds  
(without alignments)  
182.261 Million cell updates/sec

Title: US-09-900-147-7  
Perfect score: 31  
Sequence: 1 ALNVLMVA 7

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 830525 seqs, 258052604 residues

Total number of hits satisfying chosen parameters: 146963

Minimum DB seq length: 0  
Maximum DB seq length: 100

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

## Database :

SPTREMBL 23:\*\*  
1: sp\_archaea:\*  
2: sp\_bacteria:\*  
3: sp\_fungi:\*  
4: sp\_human:\*  
5: sp\_invertebrate:\*  
6: sp\_mammal:\*  
7: sp\_mhc:\*  
8: sp\_mhc:\*  
9: sp\_organelle:\*  
10: sp\_plant:\*  
11: sp\_rodent:\*  
12: sp\_virus:\*  
13: sp\_vertebrate:\*  
14: sp\_unclassified:\*  
15: sp\_rvirus:\*  
16: sp\_bacteriap:\*  
17: sp\_archaeap:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	25	80.6	53	16	Q9KCV9
2	24	77.4	50	2	Q51684
3	24	77.4	56	5	Q9UAC6
4	24	77.4	64	16	Q8NQT3
5	24	77.4	74	10	Q9ZRT2
6	24	77.4	98	8	Q21582
7	23	74.2	51	16	Q92XA1
8	23	74.2	59	5	Q9W5W3
9	23	74.2	74	17	Q9HHQ1
10	23	74.2	76	11	O70322
11	23	74.2	81	16	Q9CNL3
12	23	74.2	90	2	P72475
13	23	74.2	95	16	Q8KCM5
14	23	74.2	98	8	Q9MES8
15	23	74.2	98	8	Q9W9B6
16	23	74.2	98	8	Q21588

17	23	74.2	98	8	Q21593
18	23	74.2	98	8	Q21541
19	22	71.0	25	5	Q9BM46
20	22	71.0	25	5	Q9BM45
21	22	71.0	32	2	Q8GPX7
22	22	71.0	42	16	Q8PO15
23	22	71.0	47	9	Q9MCH4
24	22	71.0	54	2	Q52616
25	22	71.0	55	16	Q92G28
26	22	71.0	56	10	Q38984
27	22	71.0	57	13	Q9PSL0
28	22	71.0	61	4	Q96H27
29	22	71.0	64	16	Q8DV25
30	22	71.0	65	16	Q8ZF66
31	22	71.0	68	13	Q9YI43
32	22	71.0	68	16	Q8YP72
33	22	71.0	75	10	Q8RZP1
34	22	71.0	77	12	Q84212
35	22	71.0	81	2	Q9ZH41
36	22	71.0	82	12	Q98461
37	22	71.0	83	9	Q48385
38	22	71.0	86	16	Q92FW3
39	22	71.0	89	8	Q95888
40	22	71.0	90	2	Q8GC58
41	22	71.0	91	16	Q8XYQ6
42	22	71.0	91	16	Q8NL30
43	22	71.0	92	2	Q93F09
44	22	71.0	92	16	Q8FG80
45	22	71.0	95	16	Q92JC7

## ALIGNMENTS

### RESULT 1

Q9KCV9 ID Q9KCV9 PRELIMINARY; PRT; 53 AA.  
AC Q9KCV9;  
DT 01-OCT-2000 (TrEMBLrel. 15, Created)  
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)  
DT 01-MAR-2002 (TrEMBLrel. 20, Last annotation update)  
DE Hypothetical protein BH1460.  
GN BH1460.  
OS Bacillus halodurans.  
OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.  
OX NCBI\_TaxID=86665;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=C-125 / JCM 9153;  
RX MEDLINE=20512582; PubMed=11058132;  
RA Takami H., Nakasone K., Takaki Y., Maeno G., Sasaki R., Masui N., Fuji F., Hirama C., Nakamura Y., Ogasawara N., Kuhara S., RA Horikoshi K.;  
RT "Complete genome sequence of the alkaliphilic bacterium Bacillus RT halodurans and genomic sequence comparison with Bacillus subtilis."; RL Nucleic Acids Res. 28:4317-4331(2000).  
DR EMBL; AP001512; BAB05179.1;--  
KW Hypothetical protein; Complete proteome.  
SQ SEQUENCE 53 AA; 5537 MW; 2F97204707F6EE46 CRC64;

Query Match 80.6%; Score 25; DB 16; Length 53;  
Best Local Similarity 85.7%; Pred. No. 1.3e+02;  
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 ALNVLMVA 7  
|||  
Db 15 ALNVLMVA 21

### RESULT 2

Q51684 ID Q51684 PRELIMINARY; PRT; 50 AA.  
AC Q51684;

DT 01-NOV-1996 (TrEMBLrel. 01, Created)  
 DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)  
 DT 01-NOV-1998 (TrEMBLrel. 08, Last annotation update)  
 DE CcoH (Fragment).  
 GN CCOH.  
 OS Paracoccus denitrificans.  
 OC Bacteria; Proteobacteria; Alphaproteobacteria; Rhodobacterales;  
 OC Rhodobacteraceae; Paracoccus.  
 OX NCBI\_TaxID=266;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=Pd1222;  
 RA MEDLINE=96405647; PubMed=8809776;  
 RA de Gier J.W., Schepper M., Reijnders W.N.M., van Dyck S.J.,  
 RA Slotboom D.J., Warne A., Saraste M., Krab K., Finel M.,  
 RA Stouthamer A.H., van Spanning R.J.M., der Oost J.;  
 RT "Structural and functional analysis of aas-type and cbb3-type  
 RT cytochrome c oxidases of Paracoccus denitrificans reveals significant  
 RT differences in proton-pump design."  
 RL Mol. Microbiol. 20:1247-1260(1996).  
 DR EMBL; U34353; AAC44521.1; -.  
 FT NON\_TER 50  
 SQ SEQUENCE 50 AA; 5305 MW; 34C7C5FA992BDD8A CRC64;  
 Query Match 77.4%; Score 24; DB 2; Length 50;  
 Best Local Similarity 71.4%; Pred. No. 2.2e+02;  
 Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;  
 QY 1 ALNVLM 7  
 Db 23 AVNVFMA 29  
 RESULT 3  
 Q9UAC6 PRELIMINARY; PRT; 56 AA.  
 ID Q9UAC6  
 AC Q9UAC6  
 DT 01-MAY-2000 (TrEMBLrel. 13, Created)  
 DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)  
 DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)  
 DE Bmp-03 neurotoxin precursor.  
 OS Mesobuthus martensii (Machurian scorpion) (Buthus martensii).  
 OC Eukaryota; Metazoa; Arthropoda; Chelicerata; Arachnida; Scorpiones;  
 OC Buthoidea; Buthidae; Mesobuthus.  
 OX NCBI\_TaxID=34649;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Wu J.J., Dai L., Chi C.W.;  
 RA Submitted (JUL-1998) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; AF079063; RAD47377.1; -.  
 DR HSP; Q9NJ77; 1DUP.  
 KW Neurotoxin; Signal. 28  
 FT SIGNAL 1  
 FT CHAIN 29 56  
 SQ SEQUENCE 56 AA; 5940 MW; 4BB7C73200798CDF CRC64;  
 Query Match 77.4%; Score 24; DB 5; Length 56;  
 Best Local Similarity 71.4%; Pred. No. 2.4e+02;  
 Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;  
 QY 1 ALNVLM 7  
 Db 13 AMNVFMA 19  
 RESULT 4  
 Q9NOT3 PRELIMINARY; PRT; 64 AA.  
 ID Q9NOT3  
 AC Q9NOT3  
 DT 01-OCT-2002 (TrEMBLrel. 22, Created)  
 DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)  
 DT 01-OCT-2002 (TrEMBLrel. 22, Last annotation update)  
 DE Hypothetical protein Cg11341.  
 GN CGL1341.  
 OS Corynebacterium glutamicum (Brevibacterium flavum).  
 OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;  
 OC Corynebacterineae; Corynebacteriaceae; Corynebacterium.  
 OX NCBI\_TaxID=1718;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=ATCC 13032 / DSM 20300 / NCIB 10025;  
 RA Nakagawa S.;  
 RT "Complete genomic sequence of Corynebacterium glutamicum ATCC 13032.";  
 RT Submitted (MAY-2002) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; AP005279; BAB98734.1; -.  
 KW Hypothetical protein; Complete proteome.  
 SQ SEQUENCE 64 AA; 7524 MW; 2B9CBEC70EA2897 CRC64;  
 Query Match 77.4%; Score 24; DB 16; Length 64;  
 Best Local Similarity 83.3%; Pred. No. 2.8e+02;  
 Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 ALNVLM 6  
 Db 24 ALNVLL 29  
 RESULT 5  
 Q9ZRT2 PRELIMINARY; PRT; 74 AA.  
 ID Q9ZRT2  
 AC Q9ZRT2  
 DT 01-MAY-1999 (TrEMBLrel. 10, Created)  
 DT 01-MAY-1999 (TrEMBLrel. 10, Last sequence update)  
 DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)  
 DE MSH2 protein (Fragment).  
 GN MSH2.  
 OS Triticum aestivum (Wheat).  
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Pooideae;  
 OC Triticeae; Triticum.  
 OX NCBI\_TaxID=4565;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA STRAIN=cv. Chinese Spring;  
 RX PubMed=10659795;  
 RA Korzun V., Boerner A., Siebert R., Malyshev S., Hilpert M., Kunze R.,  
 RA Puchta H.;  
 RT "Chromosomal location and genetic mapping of the mismatch repair gene  
 RT homologs MSH2, MSH3 and MSH6 in rye and wheat."  
 RL Genome 42:1255-1257(1999).  
 DR EMBL; AJ131671; CAA10479.1; -.  
 DR InterPro; IPR000432; Muts\_C.  
 DR Pfam; PF00488; Muts\_C; 1.  
 DR ProDom; PD001263; Muts\_C; 1.  
 DR SMART; SM00534; MUTSAC; 1.  
 FT NON\_TER 1  
 FT NON\_TER 74 74  
 SQ SEQUENCE 74 AA; 7880 MW; 578261DB72AE93A9 CRC64;  
 Query Match 77.4%; Score 24; DB 10; Length 74;  
 Best Local Similarity 83.3%; Pred. No. 3.2e+02;  
 Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;  
 QY 2 LNVLM 7  
 Db 11 VNVLM 16  
 RESULT 6  
 Q21582 PRELIMINARY; PRT; 98 AA.  
 ID Q21582  
 AC Q21582  
 DT 01-JAN-1998 (TrEMBLrel. 05, Created)  
 DT 01-JAN-1998 (TrEMBLrel. 05, Last sequence update)  
 DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)  
 DE NADH dehydrogenase subunit 4L.

```

GN ND4L.
OS Scotinomyia teguina (Alston's brown mouse).
OG Mitochondrion.
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Sigmodontinae;
OC Scotinomyia.
OX NCBI_TaxID=56236;
RN [1]
RN SEQUENCE FROM N.A.
RP MEDLINE=98152303; PubMed=9491603;
RA Engel S.R., Hogan K.M., Taylor J.F., Davis S.K.;
RT "Molecular systematics and paleobiogeography of the South American
RT sigmodontine rodents.";
RL Mol. Biol. Evol. 15:35-49 (1998).
DR ENBL; U83828; AAB87245.1; -.
DR InterPro; IPR001133; Oxidored4L.
DR InterPro; IPR003214; Oxidored4L.
DR Pfam; PF00430; oxidored_2; 1.
DR ProDom; PD000359; Oxidred4L; 1.
KW Mitochondrion.
SQ SEQUENCE 98 AA; 10742 MW; F297223A06492F8A CRC64;

Query Match 77.4%; Score 24; DB 8; Length 98;
Best Local Similarity 66.7%; Pred. No. 4.1e+02;
Matches 4; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Qy 2 LNVLMA 7
Db 6 MNILMA 11
:::|

RESULT 7
Q92XA1 PRELIMINARY; PRT; 51 AA.
AC Q92XA1;
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-MAR-2002 (TrEMBLrel. 20, Last annotation update)
DE Hypothetical protein RB0061.
GN RB0061 OR SMB20061.
OS Rhizobium meliloti (Sinorhizobium meliloti).
OG Plasmid pSymb (megaplasmid 2).
OC Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;
OC Rhizobiaceae; Sinorhizobium.
OX NCBI_TaxID=382;
RN [1]
RN SEQUENCE FROM N.A.
RP STRAIN=1021;
RX MEDLINE=21396508; PubMed=11481431;
RA Finan T.M., Weidner S., Wong K., Buhrmester J., Chain P.,
RA Vorhoeelter F.J., Hernandez-Lucas I., Becker A., Cowie A., Gouzy J.,
RA Golding B., Puehler A.;
RT "The complete sequence of the 1,683-kb pSymb megaplasmid from the N2-
RT fixing endosymbiont Sinorhizobium meliloti.";
RL Proc. Natl. Acad. Sci. U.S.A. 98:9889-9894 (2001).
DR ENBL; AL603642; CAC48461.1; -.
KW Plasmid; Hypothetical protein; Complete proteome.
SQ SEQUENCE 51 AA; 5923 MW; OCC8242997150D7E CRC64;

Query Match 74.2%; Score 23; DB 16; Length 51;
Best Local Similarity 71.4%; Pred. No. 4e+02;
Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 1 ALNVLMA 7
Db 33 ALNLLFA 39
||:|

RESULT 8
Q9WSW3 PRELIMINARY; PRT; 59 AA.
ID Q9WSW3
AC Q9WSW3;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)

```

```

DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-MAY-2000 (TrEMBLrel. 13, Last annotation update)
DE CG14474 protein.
GN CG14474.
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
RN [1]
RN SEQUENCE FROM N.A.
RP STRAIN=BERKELEY.
RX MEDLINE=20196006; PubMed=107311132;
RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA Ananides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
RA George R.G., Lewis S.B., Richards S., Ashburner M., Henderson S.N.,
RA Sutton G.G., Wortman J.R., Vandal M.D., Zhang Q., Chen L.X.,
RA Brandon R.C., Rogers Y.-H.C., Blazej R.G., Champe M., Pfeiffer B.D.,
RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
RA Abril J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,
RA Bailew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
RA Beeson K.Y., Benos P.V., Bertram B.P., Bhandari D., Bolshakov S.,
RA Borkova D., Botchan M.R., Bouck J., Brokstein P., Brotter P.,
RA Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA Cherry J.M., Cui Y., Dahlke C., Davenport L.B., Davies P.,
RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,
RA Fessler C., Gabriellian A.E., Garg N.S., Gelbart W.M., Glasser K.,
RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,
RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
RA Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
RA Nelson D.R., Nelson K.A., Nixon K., Nuskern D.R., Pacleb J.M.,
RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
RA Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
RA Wang Z.-Y., Wassarman D.A., Weinstein G.C., Wu D., Yang S., Yao Q.A.,
RA Williams S.M., Woodage T., Worley K.C., Wu D., Zhang G., Zhao L.,
RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhou X., Zhu S., Smith H.O.,
RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Smith H.O.,
RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
RT "The genome sequence of Drosophila melanogaster.";
RL Science 287:2185-2195 (2000).
DR ENBL; AE002612; AAP45482.1; -.
DR FlyBase; FBgn0040055; CG14474.
SQ SEQUENCE 59 AA; 6824 MW; 37EB14DF0D0744E0 CRC64;

Query Match 74.2%; Score 23; DB 5; Length 59;
Best Local Similarity 83.3%; Pred. No. 4.5e+02;
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 2 LNVLMA 7
Db 39 LNALMA 44
|||

RESULT 9
Q9HHQ1 PRELIMINARY; PRT; 74 AA.
ID Q9HHQ1
AC Q9HHQ1;
DT 01-MAR-2001 (TrEMBLrel. 16, Created)
DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT 01-MAR-2001 (TrEMBLrel. 16, Last annotation update)
DE Vng6287H.
GN VNG6287H.

```

Tue Feb 17 11:55:57 2004

us-09-900-147-7.rsp

```

OS Halobacterium sp. (strain NRC-1).
OG Plasmid pNRC200.
OC Archaea; Euryarchaeota; Halobacteriia; Halobacteriales;
OC Halobacteriaceae; Halobacterium.
OX NCBI_TaxID=64091;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=20504483; PubMed=11016950;
RA Ng W.V., Kennedy S.P., Mahairas G.G., Berquist B., Pan M.,
RA Shukla H.D., Laaky S.R., Baliga N.S., Thorsen V., Sbrogna J.,
RA Swartzell S., Weir D., Hall J., Dahl T.A., Welter R., Goo Y.A.,
RA Leithausen D.G., Keller K., Cruz R., Danson M.J., Hough D.W.,
RA Maddocks D.G., Beck R.F., Pohlschroder M., Spudich J.L., Jung K.-H.,
RA Isenbarger T.A., Peck R.F., Daniels C.J., Dennis P.P., Omer A.D.,
RA Alam M., Freitas T., Hou S., Liang P., Riley M., Hood L., DasSarma S.;
RA Ehardt H., Lowe T.M., Liang P., Riley M., Hood L., DasSarma S.;
RT "Genome sequence of Halobacterium species NRC-1";
RT Proc. Natl. Acad. Sci. U.S.A. 97:12176-12181(2000).
RL EMBL; AF045159; AAC20925.1; -.
KW Plasmid; Complete proteome.
SQ SEQUENCE 74 AA; 8286 MW; 5D045F9B606B6FC5 CRC64;

Query Match 74.2%; Score 23; DB 17; Length 74;
Best Local Similarity 83.3%; Pred. No. 5.6e+02;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 2 LNVLMA 7
| | | | |
DB 47 LNVLVA 52

RESULT 10
O70322 PRELIMINARY; PRT; 76 AA.
AC O70322;
DT 01-AUG-1998 (TRENBLrel. 07, Created)
DT 01-AUG-1998 (TRENBLrel. 07, Last sequence update)
DT 01-OCT-2002 (TRENBLrel. 22, Last annotation update)
DE ABC transporter (Fragment).
GN ABCB6 OR 1200005B17RIK.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RA Stuart R.O., Pavlova A., Nigam S.K.;
RA Submitted (JAN-1998) to the EMBL/GenBank/DBJ databases.
RL EMBL; AF045562; AAC05261.1; -.
DR MGD; MGI:1921354; Abcb6.
DR NON TER 76
FT NON TER 76
SQ SEQUENCE 76 AA; 8600 MW; 7D614DBE4333E26A CRC64;

Query Match 74.2%; Score 23; DB 11; Length 76;
Best Local Similarity 83.3%; Pred. No. 5.8e+02;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 ALNVLV 6
| | | | |
DB 6 ALNVLV 11

RESULT 11
Q9CNL3 PRELIMINARY; PRT; 81 AA.
AC Q9CNL3;
DT 01-JUN-2001 (TRENBLrel. 17, Created)
DT 01-JUN-2001 (TRENBLrel. 17, Last sequence update)
DT 01-DEC-2001 (TRENBLrel. 19, Last annotation update)
DE Hypothetical protein PM0415.
GN PM0415.
OS Pasteurella multocida.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Pasteurellales;

OC Pasteurellaceae; Pasteurella.
OX NCBI_TaxID=747;
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN=Pm70;
RA May B.J., Zhang Q., Li L.L., Paustian M.L., Whittam T.S., Kapur V.;
RA "Complete genomic sequence of Pasteurella multocida Pm70.";
RA Proc. Natl. Acad. Sci. U.S.A. 98:3460-3465(2001).
RL EMBL; AE006077; AA02499.1; -.
KW Hypothetical protein; Complete proteome.
SQ SEQUENCE 81 AA; 9432 MW; 0FB82498E3551973 CRC64;

Query Match 74.2%; Score 23; DB 16; Length 81;
Best Local Similarity 57.1%; Pred. No. 6.1e+02;
Matches 4; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 1 ALNVLMA 7
| | | | |
DB 31 ALNVLMS 37

RESULT 12
P72475 PRELIMINARY; PRT; 90 AA.
AC P72475;
DT 01-FEB-1997 (TRENBLrel. 02, Created)
DT 01-FEB-1997 (TRENBLrel. 02, Last sequence update)
DT 01-MAR-2003 (TRENBLrel. 23, Last annotation update)
DE High affinity branched chain amino acid transport protein (Fragment).
DE (Fragment).
GN LIVG.
OS Streptococcus mutans.
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC Streptococcus.
OX NCBI_TaxID=1309;
RN [1]
RP SEQUENCE FROM N.A.
RA Peruzzi F., Piggot P.J., Daneo-Moore L.;
RA Submitted (OCT-1996) to the EMBL/GenBank/DBJ databases.
RL EMBL; U75471; AAB41191.1; -.
DR InterPro; IPR003439; ABC transporter.
DR ProDom; PD000006; ABC transporter; 1.
DR PROSITE; PS00211; ABC_TRANSPORTER; 1.
FT NON TER 1
FT NON TER 90
FT NON TER 90
SQ SEQUENCE 90 AA; 9857 MW; 3F92DE952FE38647 CRC64;

Query Match 74.2%; Score 23; DB 2; Length 90;
Best Local Similarity 71.4%; Pred. No. 6.8e+02;
Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 ALNVLMA 7
| | | | |
DB 67 ALSILMA 73

RESULT 13
Q8KCM5 PRELIMINARY; PRT; 95 AA.
AC Q8KCM5;
DT 01-OCT-2002 (TRENBLrel. 22, Created)
DT 01-OCT-2002 (TRENBLrel. 22, Last sequence update)
DT 01-OCT-2002 (TRENBLrel. 22, Last annotation update)
DE Hypothetical protein CTI389.
GN CTI389.
OS Chlorobium tepidum.
OC Bacteria; Chlorobi; Chlorobia; Chlorobiales; Chlorobiaceae;
OC Chlorobium.
OX NCBI_TaxID=1097;
RN [1]
RP SEQUENCE FROM N.A.

```

```

RC STRAIN-TLS / ATCC 49652 / DSM 12025;
RX MEDLINE=22103685; PubMed=12093901;
RA Eisen J.A., Nelson K.E., Paulsen I.T., Heidelberg J.P., Wu M.,
RA Dodson R.J., Deboy R., Gwinn M.L., Nelson W.C., Haft D.H.,
RA Hickey E.K., Peterson J.D., Durkin A.S., Kolonay J.L., Yang F.,
RA Holt I., Umayam L.A., Mason T., Brenner M., Shea T.P., Parksey D.,
RA Nieman W.C., Feldblyum T.V., Hansen C.L., Craven M.B., Radune D.,
RA Vamathevan J., Khouri H., White O., Gruber T.M., Ketchum K.A.,
RA Venter J.C., Tettelin H., Bryant D.A., Fraser C.M.;
RT "The complete genome sequence of Chlorobium tepidum TLS, a
RL photosynthetic, anaerobic, green-sulfur bacterium.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:9509-9514(2002).
DR EMBL; AF012897; AAM72618.1; -.
DR TIGR; CT1389; -.
KW Hypothetical protein; Complete proteome.
SQ SEQUENCE 95 AA; 10702 MW; 529E865CCEB2816F CRC64;

Query Match 74.2%; Score 23; DB 16; Length 95;
Best Local Similarity 66.7%; Pred. No. 7.1e+02;
Matches 4; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ALNVLM 6
Db 50 ALNILL 55

RESULT 14
Q9MES8 QMES8 PRELIMINARY; PRT; 98 AA.
AC QMES8;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DE 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE NADH dehydrogenase subunit 4L.
GN NADH4L.
OS Physeter catodon (Sperm whale) (Physeter macrocephalus).
OG Mitochondrion.
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Cetacea; Odontoceti;
OC Physeteridae; Physeter.
OX NCBI_TaxID=9755;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=20296818; PubMed=10835487;
RA Arnason U., Gullberg A., Gretaarsdottir S., Ursing B., Janke A.;
RT "The mitochondrial genome of the sperm whale and a new molecular
RT reference for estimating eutherian divergence dates.";
RL J. Mol. Evol. 50:569-578(2000).
DR EMBL; AJ277029; CAB98278.1; -.
DR InterPro; IPR001133; Oxidored4L.
DR InterPro; IPR003214; Oxidred4L.
DR Pfam; PF00420; oxidored g2; 1.
DR ProDom; PD000359; Oxidred4L; 1.
KW Mitochondrion.
SQ SEQUENCE 98 AA; 10767 MW; B2BB09927D25AD1B CRC64;

Query Match 74.2%; Score 23; DB 8; Length 98;
Best Local Similarity 66.7%; Pred. No. 7.3e+02;
Matches 4; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Qy 2 LNVLMA 7
Db 6 MNVWMA 11

RESULT 15
Q8W9B6 Q8W9B6 PRELIMINARY; PRT; 98 AA.
AC Q8W9B6;
DT 01-MAR-2002 (TrEMBLrel. 20, Created)
DT 01-MAR-2002 (TrEMBLrel. 20, Last sequence update)
DE 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE NADH dehydrogenase subunit 4L.

```

```

GN NADH4L.
OS Vombatus ursinus (Common wombat).
OG Mitochondrion.
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Metatheria; Diprotodontia; Vombatidae; Vombatus.
OX NCBI_TaxID=29139;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=21592585; PubMed=11734900;
RA Janke A., Magnell O., Wiczyrek G., Westerman M., Arnason U.;
RT "Phylogenetic analysis of 18S rRNA and the mitochondrial genomes of
RT the wombat, Vombatus ursinus, and the spiny anteater, Tachyglossus
RT aculeatus: increased support for the Marsupionta Hypothesis.";
RL J. Mol. Evol. 54:71-80(2002).
DR EMBL; AJ304826; CAC83113.1; -.
DR InterPro; IPR001133; Oxidored4L.
DR InterPro; IPR003214; Oxidred4L.
DR Pfam; PF00420; oxidored g2; 1.
DR ProDom; PD000359; Oxidred4L; 1.
KW Mitochondrion.
SQ SEQUENCE 98 AA; 10655 MW; E15DFB15EC4219CA CRC64;

Query Match 74.2%; Score 23; DB 8; Length 98;
Best Local Similarity 57.1%; Pred. No. 7.3e+02;
Matches 4; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ALNVLMA 7
Db 5 SLNLIMA 11

RESULT 16
O21588 O21588 PRELIMINARY; PRT; 98 AA.
AC O21588;
DT 01-JAN-1998 (TrEMBLrel. 05, Created)
DT 01-JAN-1998 (TrEMBLrel. 05, Last sequence update)
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE NADH dehydrogenase subunit 4L.
GN N04L.
OS Ochrotomys nuttali (Golden mouse) (Peromyscus nuttali).
OG Mitochondrion.
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Sigmodontinae;
OC Ochrotomys.
OX NCBI_TaxID=56229;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=98152303; PubMed=9491603;
RA Engel S.R., Hogan K.M., Taylor J.F., Davis S.K.;
RT "Molecular systematics and paleobiogeography of the South American
RT sigmodontine rodents.";
RL Mol. Biol. Evol. 15:35-49(1998).
DR EMBL; U83830; AAB87209.1; -.
DR InterPro; IPR001133; Oxidored4L.
DR InterPro; IPR003214; Oxidred4L.
DR Pfam; PF00420; oxidored g2; 1.
DR ProDom; PD000359; Oxidred4L; 1.
KW Mitochondrion.
SQ SEQUENCE 98 AA; 10728 MW; DAE4B4DE0EA713C CRC64;

Query Match 74.2%; Score 23; DB 8; Length 98;
Best Local Similarity 66.7%; Pred. No. 7.3e+02;
Matches 4; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Qy 2 LNVLMA 7
Db 6 LNVLLA 11

RESULT 17
O21593 O21593 PRELIMINARY; PRT; 98 AA.

```

AC 021593;  
 DT 01-JAN-1998 (T-EMBLrel. 05, Created)  
 DT 01-JAN-1998 (T-EMBLrel. 05, Last sequence update)  
 DT 01-JUN-2002 (T-EMBLrel. 21, Last annotation update)  
 DE NADH dehydrogenase subunit 4L.  
 GN ND4L.  
 OS Reithodontomys fulvescens (Fulvous harvest mouse).  
 OS Mitochondrion.  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Sigmodontinae;  
 OC Reithodontomys.  
 OC NCBI\_TaxID=56213;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=98152303; PubMed=9491603;  
 RA Engel S.R., Hogan K.M., Taylor J.F., Davis S.K.;  
 RT "Molecular systematics and paleobiogeography of the South American  
 RL Mol. Biol. Evol. 15:35-49 (1998).  
 DR EMBL; U83832; AAB87233.1; -.  
 DR InterPro; IPR001133; Oxidored4L.  
 DR InterPro; IPR003214; Oxidred4L.  
 DR Pfam; PF00420; oxidored\_Q2; 1.  
 DR ProDom; PD000359; Oxidred4L; 1.  
 KW Mitochondrion.  
 SQ SEQUENCE 98 AA; 10774 MW; 653067658DBD77EC CRC64;

Query Match 74.2%; Score 23; DB 8; Length 98;  
 Best Local Similarity 66.7%; Pred. No. 7.3e+02;  
 Matches 4; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 2 LNVLMA 7  
 Db 6 LNILLA 11

RESULT 18  
 Q21541 PRELIMINARY; PRT; 98 AA.  
 AC 021541;  
 DT 01-JAN-1998 (T-EMBLrel. 05, Created)  
 DT 01-JAN-1998 (T-EMBLrel. 05, Last sequence update)  
 DT 01-JUN-2002 (T-EMBLrel. 21, Last annotation update)  
 DE NADH dehydrogenase subunit 4L.  
 GN ND4L.  
 OS Bolomys lactens (rufous-bellied bolo mouse).  
 OS Mitochondrion.  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Sigmodontinae;  
 OC Bolomys.  
 OC NCBI\_TaxID=56217;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=98152303; PubMed=9491603;  
 RA Engel S.R., Hogan K.M., Taylor J.F., Davis S.K.;  
 RT "Molecular systematics and paleobiogeography of the South American  
 RL sigmodontine rodents";  
 RL Mol. Biol. Evol. 15:35-49 (1998).  
 DR EMBL; U83813; AAB87158.1; -.  
 DR InterPro; IPR001133; Oxidored4L.  
 DR InterPro; IPR003214; Oxidred4L.  
 DR Pfam; PF00420; oxidored\_Q2; 1.  
 DR ProDom; PD000359; Oxidred4L; 1.  
 KW Mitochondrion.  
 SQ SEQUENCE 98 AA; 10785 MW; BD00880FEE64CDF4 CRC64;

Query Match 74.2%; Score 23; DB 8; Length 98;  
 Best Local Similarity 66.7%; Pred. No. 7.3e+02;  
 Matches 4; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 2 LNVLMA 7  
 Db 6 LNILLA 11

RESULT 19  
 Q9BM46 PRELIMINARY; PRT; 25 AA.  
 ID Q9BM46;  
 AC Q9BM46;  
 DT 01-JUN-2001 (T-EMBLrel. 17, Created)  
 DT 01-JUN-2001 (T-EMBLrel. 17, Last sequence update)  
 DT 01-JUN-2001 (T-EMBLrel. 17, Last annotation update)  
 DE LINE-like reverse transcriptase (Fragment).  
 OS Euperipatoides rowelli.  
 OC Eukaryota; Metazoa; Onychophora; Peripatopsidae; Euperipatoides.  
 OC NCBI\_TaxID=49087;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TRANSPOSON=LRT-L2 retrotransposon;  
 RX MEDLINE=20570504; PubMed=11121049;  
 RA Arkhipova I., Meselson M.;  
 RT "Transposable elements in sexual and ancient asexual taxa.";  
 RL Proc. Natl. Acad. Sci. U.S.A. 97:14473-14477 (2000).  
 DR EMBL; AV013949; AAG59932.1; -.  
 DR RNA-directed DNA polymerase.  
 FT NON\_TER 1 1  
 FT NON\_TER 25 25  
 SQ SEQUENCE 25 AA; 2809 MW; 5BBAE0AE0458F32 CRC64;

Query Match 71.0%; Score 22; DB 5; Length 25;  
 Best Local Similarity 80.0%; Pred. No. 3.6e+02;  
 Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 3 NVLMA 7  
 Db 9 NILMA 13

RESULT 20  
 Q9BM45 PRELIMINARY; PRT; 25 AA.  
 ID Q9BM45;  
 AC Q9BM45;  
 DT 01-JUN-2001 (T-EMBLrel. 17, Created)  
 DT 01-JUN-2001 (T-EMBLrel. 17, Last sequence update)  
 DT 01-JUN-2001 (T-EMBLrel. 17, Last annotation update)  
 DE LINE-like reverse transcriptase (Fragment).  
 OS Euperipatoides rowelli.  
 OC Eukaryota; Metazoa; Onychophora; Peripatopsidae; Euperipatoides.  
 OC NCBI\_TaxID=49087;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TRANSPOSON=LRT-L3 retrotransposon;  
 RX MEDLINE=20570504; PubMed=11121049;  
 RA Arkhipova I., Meselson M.;  
 RT "Transposable elements in sexual and ancient asexual taxa.";  
 RL Proc. Natl. Acad. Sci. U.S.A. 97:14473-14477 (2000).  
 DR EMBL; AV013950; AAG59933.1; -.  
 DR RNA-directed DNA polymerase.  
 FT NON\_TER 1 1  
 FT NON\_TER 25 25  
 SQ SEQUENCE 25 AA; 2757 MW; ED30E0AD8052B998 CRC64;

Query Match 71.0%; Score 22; DB 5; Length 25;  
 Best Local Similarity 80.0%; Pred. No. 3.6e+02;  
 Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 3 NVLMA 7  
 Db 9 NILMA 13

RESULT 21  
 Q8GPX7 PRELIMINARY; PRT; 32 AA.  
 ID Q8GPX7  
 AC Q8GPX7;  
 DT 01-MAR-2003 (T-EMBLrel. 23, Created)

```

DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE Hypothetical protein.
OC Pseudomonas aeruginosa.
OS Bacteria; Proteobacteria; Gammaproteobacteria; Pseudomonadales;
OC Pseudomonadaceae; Pseudomonas.
OX NCBI_TaxID=287;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=SG17M;
RX MEDLINE=22313472; PubMed=12426355;
RA Larbig K.D., Christmann A., Johann A., Klockgether J., Hartsch T.,
RA Merkl R., Wiehmann L., Fritz H.J., Tummeler B.;
RT "Gene Islands Integrated into tRNA(Gly) Genes Confer Genome Diversity
RT on a Pseudomonas aeruginosa Clone.";
RL J. Bacteriol. 184:6665-6680(2002).
DR EMBL; AF40524; AAN62247.1; -.
KW Hypothetical protein.
SQ SEQUENCE 32 AA; 3435 MW; 6281AD17052C2682 CRC64;

Query Match 71.0%; Score 22; DB 2; Length 32;
Best Local Similarity 100.0%; Pred. No. 4.5e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ALNVL 5
DB 26 ALNVL 30

RESULT 22
ID Q8P015 PRELIMINARY; PRT; 42 AA.
AC Q8P015;
DT 01-OCT-2002 (TrEMBLrel. 22, Created)
DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
DT 01-OCT-2002 (TrEMBLrel. 22, Last annotation update)
DE Hypothetical protein spyM18_1343.
GN SPYM18_1343.
OS Streptococcus pyogenes (serotype M18).
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC Streptococcus.
OX NCBI_TaxID=186103;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=MGAS8232 / Serotype M18;
RX MEDLINE=21927593; PubMed=11917108;
RA Smoot J.C., Barbican K.D., Van Gompel J.J., Smoot L.M., Chaussee M.S.,
RA Sylvia G.L., Sturdevant D.E., Rickiers S.M., Porcella S.F.,
RA Parkins L.D., Beres S.B., Campbell D.S., Smith T.M., Zhang Q.,
RA Kapur V., Daly J.A., Veasy L.G., Musser J.M.;
RT "Genome sequence and comparative microarray analysis of serotype M18
RT group A Streptococcus strains associated with acute rheumatic fever
RT outbreaks.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:4668-4673(2002).
DR EMBL; AE010055; AAL97944.1; -.
KW Hypothetical protein; Complete proteome.
SQ SEQUENCE 42 AA; 4452 MW; 48CD6FEB4B0CD799 CRC64;

Query Match 71.0%; Score 22; DB 16; Length 42;
Best Local Similarity 83.3%; Pred. No. 5.9e+02;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 ALNVL 6
DB 37 ALNVL 42

RESULT 23
Q9MCH4 PRELIMINARY; PRT; 47 AA.
AC Q9MCH4;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)

```

```

DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE Gp47.
OS Streptococcus thermophilus bacteriophage Sfi18.
OC Viruses; dsDNA viruses, no RNA stage; Caudovirales; Siphoviridae.
OX NCBI_TaxID=74382;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=99412383; PubMed=10482618;
RA Lucchini S., Desiere F., Brussow H.;
RT "Comparative genomics of Streptococcus thermophilus phage species
RT supports a modular evolution theory.";
RL J. Virol. 73:8647-8656(1999).
DR EMBL; AF158601; AAF63076.1; -.
SQ SEQUENCE 47 AA; 5650 MW; 0A50F8A33EB231C6 CRC64;

Query Match 71.0%; Score 22; DB 9; Length 47;
Best Local Similarity 71.4%; Pred. No. 6.5e+02;
Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 ALNVLMA 7
DB 28 ALNVLMA 34

RESULT 24
Q52616 PRELIMINARY; PRT; 54 AA.
AC Q52616;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE HlyC gene 5' region (fragment).
OS Proteus vulgaris.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Proteus.
OX NCBI_TaxID=585;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=89127151; PubMed=3065612;
RA Koronakis V., Hughes C.;
RT Identification of the promoters directing in vivo expression of
RT hemolysin genes in Proteus vulgaris and Escherichia coli.";
RL Mol. Gen. Genet. 213:99-104(1988).
DR EMBL; X12571; CAA31083.1; -.
DR InterPro; IPR003996; RtxC.
DR Pfam; PF02794; HlyC; 1.
FT NON_TER 54
SQ SEQUENCE 54 AA; 6205 MW; 7C9E7EF903D954AA CRC64;

Query Match 71.0%; Score 22; DB 2; Length 54;
Best Local Similarity 71.4%; Pred. No. 7.4e+02;
Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 ALNVLMA 7
DB 33 ALNVLMA 39

RESULT 25
Q92G28 PRELIMINARY; PRT; 55 AA.
AC Q92G28;
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE Similarity to proline/betaine transporter.
GN RC1297.
OS Rickettsia conorii.
OC Bacteria; Proteobacteria; Alphaproteobacteria; Rickettsiales;
OC Rickettsiaceae; Rickettsia.
OX NCBI_TaxID=781;
RN [1]
RP SEQUENCE FROM N.A.

```

Tue Feb 17 11:55:57 2004

us-09-900-147-7.rspt

```

RC STRAIN=Malish 7; PubMed=11557893;
RX MEDLINE=21442074; Reneato-Audiffren P., Fournier P.-E., Barbe V.,
RA Ogata H., Audic S., Reneato-Audiffren P., Fournier P.-E., Barbe V.,
RA Samson D., Roux V., Cossart P., Weissenbach J., Claverie J.-M.,
RA Raoult D.;
RT "Mechanisms of evolution in Rickettsia conorii and R. prowazekii.";
RL Science 293:2093-2098(2001).
DR EMBL; AE008676; AAL03835.1; -.
KW Complete proteome.
SQ SEQUENCE 55 AA; 6227 MW; E87478DCDC00174A CRC64;

Query Match 71.0%; Score 22; DB 16; Length 55;
Best Local Similarity 71.4%; Pred. No. 7.6e+02;
Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 ALNVLMA 7
Db 4 ALNVMPA 10

RESULT 26
Q38984 PRELIMINARY; PRT; 56 AA.
AC Q38984; 1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE Protein kinase catalytic domain (Fragment).
GN AK19.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eumossids II; Brassicales; Brassicaceae; Arabidopsids.
OX NCBI_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=El1-O; TISSUE=Leaf;
RX MEDLINE=96123233; PubMed=8534852;
RA Thummler F., Kirchner M., Teuber R., Dittrich P.;
RT "Differential accumulation of the transcripts of 22 novel protein
RT kinase genes in Arabidopsis thaliana.";
RL Plant Mol. Biol. 29:551-565(1995).
DR EMBL; X86964; CAA60527.1; -.
DR InterPro; IPR000719; Prot_kinase.
DR Pfam; PF00069; pkinase; 1.
DR ProDom; PD000001; Prot_kinase; 1.
DR PROSITE; PS50011; PROTEIN_KINASE_DOM; 1.
KW ATP-binding; Kinase; Transferase.
FT NON_TER 1
FT SEQUENCE 56 AA; 5804 MW; 59123279AED5D111 CRC64;

Query Match 71.0%; Score 22; DB 10; Length 56;
Best Local Similarity 80.0%; Pred. No. 7.7e+02;
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 3 NVLMA 7
Db 4 NILMA 8

RESULT 27
Q9PSL0 PRELIMINARY; PRT; 57 AA.
AC Q9PSL0; 2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE Protein kinase (Fragment).
OS Gallus sp.
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
OC Gallus.
OX NCBI_TaxID=9036;

```

```

RN SEQUENCE FROM N.A.
RP MEDLINE=93096482; PubMed=1281306;
RX Marcelle C., Eichmann A.;
RA "Molecular cloning of a family of protein kinase genes expressed in
RT the avian embryo.";
RL Oncogene 7:2479-2487(1992).
DR InterPro; IPR000719; Prot_kinase.
DR Pfam; PF00069; pkinase; 1.
DR ProDom; PD000001; Prot_kinase; 1.
DR PROSITE; PS50011; PROTEIN_KINASE_DOM; 1.
KW ATP-binding; Transferase.
FT NON_TER 1
FT NON_TER 57
FT SEQUENCE 57 AA; 6347 MW; 4F96EA9245FEAA86 CRC64;

Query Match 71.0%; Score 22; DB 13; Length 57;
Best Local Similarity 71.4%; Pred. No. 7.8e+02;
Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 ALNVLMA 7
Db 2 ARNVLLA 8

RESULT 28
Q96HZ7 PRELIMINARY; PRT; 61 AA.
AC Q96HZ7; 2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-OCT-2002 (TrEMBLrel. 22, Last annotation update)
DE Hypothetical protein.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC Tissue=Lung;
RA Strausberg R.;
RL Submitted (MAY-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC007928; AAH07928.1; -.
KW Hypothetical protein.
SQ SEQUENCE 61 AA; 6513 MW; 03E51D2F5484DBAA CRC64;

Query Match 71.0%; Score 22; DB 4; Length 61;
Best Local Similarity 71.4%; Pred. No. 8.3e+02;
Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 ALNVLMA 7
Db 29 ALRVLLA 35

RESULT 29
Q8DV25 PRELIMINARY; PRT; 64 AA.
AC Q8DV25; 2003 (TrEMBLrel. 23, Created)
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DE Putative ferredoxin.
GN SMU.694C.
OS Streptococcus mutans.
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC Streptococcus.
OX NCBI_TaxID=1309;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=UAI59 / ATCC 700610 / Serotype C;
RX MEDLINE=22295063; PubMed=12397186;
RA Ajdic D., McShan W.M., McLaughlin R.E., Savic G., Chang J.,

```



```

RA Carson M.B., Primeaux C., Tian R., Kenton S., Jia H., Lin S., Qian Y.,
RA Li S., Zhu H., Najjar F., Lai H., White J., Roe B.A., Ferretti J.J.;
RT "Genome sequence of Streptococcus mutans UA159, a cariogenic dental
RT pathogen.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:14434-14439(2002).
DR EMBL; AE014913; AAN58426.1; -.
KW Complete proteome.
SQ SEQUENCE 64 AA; 6965 MW; 424FE786CE938873 CRC64;

Query Match 71.0%; Score 22; DB 16; Length 64;
Best Local Similarity 80.0%; Pred. No. 8.7e+02;
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 3 NVLMA 7
Db 48 NILMA 52

RESULT 30
Q8ZF66 PRELIMINARY; PRT; 65 AA.
AC Q8ZF66;
DT 01-MAR-2002 (T-EMBLrel. 20, Created)
DT 01-MAR-2002 (T-EMBLrel. 20, Last sequence update)
DT 01-MAR-2003 (T-EMBLrel. 23, Last annotation update)
DE Putative membrane protein (Hypothetical).
GN YPO1852 OR Y2454.
OS Yersinia pestis.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Yersinia.
OX NCBI_TaxID=632;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CO-92 / Biovar Orientalis;
RX MEDLINE=21470413; PubMed=11586360;
RA Parkhill J., Wren B.W., Thomson N.R., Titball R.W., Holden M.T.G.,
RA Prentice M.B., Sebaiha M., James K.D., Churcher C., Mungall K.L.,
RA Baker S., Basham D., Bentley S.D., Brooks K., Cerdeno-Tarraga A.M.,
RA Chillingworth T., Cronin A., Davies R.M., Davis P., Dougan G.,
RA Feltwell T., Hamlin N., Holroyd S., Jagels K., Karlyshev A.V.,
RA Leather S., Moule S., Oyston P.C.F., Quail M., Rutherford K.,
RA Simmonds M., Skelton J., Stevens K., Whitehead S., Barrall B.G.;
RT "Genome sequence of Yersinia pestis, the causative agent of plague.";
RL Nature 413:523-527(2001).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=KIM5 / Biovar Mediaevalis;
RX MEDLINE=22137863; PubMed=12142430;
RA Deng W., Burland V., Plunkett G. III, Boutin A., Mayhew G.F., Liss P.,
RA Perna N.T., Rose D.J., Mau B., Zhou S., Schwartz D.C.,
RA Fetherston J.D., Lindler L.E., Brubaker R.R., Plano G.V.,
RA Straley S.C., McDonough K.A., Nilles M.L., Matson J.S., Blattner F.R.,
RA Parry R.D.;
RA "Genome sequence of Yersinia pestis KIM.";
RL J. Bacteriol. 194:4601-4611(2002).
DR EMBL; AJ414150; CAC90669.1; -.
DR EMBL; AE013849; AAM86011.1; -.
KW Hypothetical protein; Complete proteome.
SQ SEQUENCE 65 AA; 7315 MW; B0796B51977FBEEC CRC64;

Query Match 71.0%; Score 22; DB 16; Length 65;
Best Local Similarity 100.0%; Pred. No. 8.8e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ALNVL 5
Db 49 ALNVL 53

RESULT 31
Q9YI43 PRELIMINARY; PRT; 68 AA.
ID Q9YI43
AC Q9YI43;

```

```

DT 01-MAY-1999 (T-EMBLrel. 10, Created)
DT 01-MAY-1999 (T-EMBLrel. 10, Last sequence update)
DT 01-MAR-2003 (T-EMBLrel. 23, Last annotation update)
DE Tyrosine kinase PDGFR (Fragment).
GN PDGFR.
OS Gallus gallus (Chicken).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Archosauria; Aves; Neognathae; Galliformes; Phasianinae;
OC Gallus.
OX NCBI_TaxID=9031;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Hematopoietic bone marrow;
RA Kortschoner N., Bartunek P., Knespel S., Zenke M.;
RT "Assessing the gene expression by domain-directed differential
RT display.";
RL Submitted (DEC-1997) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF041797; AAD02125.1; -.
DR HSP; P11362; IFGK.
DR InterPro; IPR000719; Prot_kinase.
DR Pfam; PF00069; pkinase; 1.
DR ProDom; PD000001; Prot_kinase; 1.
DR PROSITE; PS50011; PROTEIN_KINASE_DOM; 1.
KW ATP-binding; Kinase; Transferase.
FT NON_TER 1
FT NON_TER 68
SQ SEQUENCE 68 AA; 7660 MW; 07C70FC6E2DC88 CRC64;

Query Match 71.0%; Score 22; DB 13; Length 68;
Best Local Similarity 71.4%; Pred. No. 9.2e+02;
Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 ALNVLMA 7
Db 7 ARNVLLA 13

RESULT 32
Q8YF72 PRELIMINARY; PRT; 68 AA.
AC Q8YF72;
DT 01-MAR-2002 (T-EMBLrel. 20, Created)
DT 01-MAR-2002 (T-EMBLrel. 20, Last sequence update)
DT 01-MAR-2002 (T-EMBLrel. 20, Last annotation update)
DE Hypothetical protein Asl4328.
GN ASL4328.
OS Anabaena sp. (strain PCC 7120).
OC Bacteria; Cyanobacteria; Nostocales; Nostocaceae; Nostoc.
OX NCBI_TaxID=103690;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=21595285; PubMed=11759840;
RA Kaneko T., Nakamura Y., Wolk C.P., Kuritz T., Sasamoto S.,
RA Watanabe A., Iriguchi M., Ishikawa A., Kawashima K., Kimura T.,
RA Kishida Y., Kohara M., Matsumoto M., Matsuno A., Muraki A.,
RA Nakazaki N., Shimoto S., Sugimoto M., Takazawa M., Yamada M.,
RA Yasuda M., Tabata S.;
RT "Complete genomic sequence of the filamentous nitrogen-fixing
RT cyanobacterium Anabaena sp. strain PCC 7120.";
RL DNA Res. 8:205-213(2001).
DR EMBL; AP003596; BAB76027.1; -.
KW Hypothetical protein; Complete proteome.
SQ SEQUENCE 68 AA; 7648 MW; 8876D5D9FFCC4B14 CRC64;

Query Match 71.0%; Score 22; DB 16; Length 68;
Best Local Similarity 71.4%; Pred. No. 9.2e+02;
Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 ALNVLMA 7
Db 45 ALELLMA 51

```

```
RESULT 33
Q8RZP1 PRELIMINARY; PRT; 75 AA.
ID Q8RZP1
AC Q8RZP1
DT 01-JUN-2002 (TrEMBLrel. 21, Created)
DT 01-JUN-2002 (TrEMBLrel. 21, Last sequence update)
DT 01-OCT-2002 (TrEMBLrel. 22, Last annotation update)
DE B1065E10.24 protein.
GN B1065E10.24
OS Oryza sativa (japonica cultivar-group).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC Ehrhartoideae; Oryzaceae; Oryza.
OX NCBI_TaxID=39947;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=cv. Nipponbare;
RA Sasaki T., Matsumoto T., Yamamoto K.;
RT "Oryza sativa nipponbare (GA3) genomic DNA, chromosome 1, BAC
clone:B1065E10.";
RL Submitted (MAY-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AP003561; BAB90373.1; -.
DR Gramene; Q8RZP1; -.
SQ SEQUENCE 75 AA; 8090 MW; 62DE4A3CF83365D6 CRC64;

Query Match 71.0%; Score 22; DB 10; Length 75;
Best Local Similarity 83.3%; Pred. No. 1e+03;
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 ALNVLN 6
Db |||||
52 AANVLN 57

RESULT 34
Q84212 PRELIMINARY; PRT; 77 AA.
ID Q84212
AC Q84212;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE ORF X protein.
OS Human papillomavirus type 41.
OC Viruses; dsDNA viruses, no RNA stage; Papillomaviridae;
OC Papillomavirus.
OX NCBI_TaxID=10589;
RN [1]
RP SEQUENCE FROM N.A.
RA Hirt L., Hirsch-Behnam A., de Villiers E.M.;
RT "Nucleotide sequence of human papillomavirus (HPV) type 41: an unusual
HPV type without a typical E2 binding site consensus sequence.";
RL Virus Res. 18:179-190(1990).
DR EMBL; X56147; CAA39620.1; -.
SQ SEQUENCE 77 AA; 9513 MW; 618D681AACDE9062 CRC64;

Query Match 71.0%; Score 22; DB 12; Length 77;
Best Local Similarity 83.3%; Pred. No. 1e+03;
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 LNVLMA 7
Db |||||
28 LNVAMA 33

RESULT 35
Q9ZH41 PRELIMINARY; PRT; 81 AA.
ID Q9ZH41
AC Q9ZH41;
DT 01-MAY-1999 (TrEMBLrel. 10, Created)
DT 01-MAY-1999 (TrEMBLrel. 10, Last sequence update)
DT 01-MAY-1999 (TrEMBLrel. 10, Last annotation update)
DE Hypothetical 8.9 kDa protein.
OS Morganella morganii (Proteus morganii).
```

```
OG Plasmid R485.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Morganella.
OX NCBI_TaxID=582;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=99047595; PubMed=9829958;
RA Hayes F.;
RT "A family of stability determinants in pathogenic bacteria.";
RL J. Bacteriol. 180:6415-6418(1998).
DR EMBL; AF072136; AAC84005.1; -.
KW Hypothetical protein; Plasmid.
SQ SEQUENCE 81 AA; 8925 MW; E42EB1330E90AD41 CRC64;

Query Match 71.0%; Score 22; DB 2; Length 81;
Best Local Similarity 66.7%; Pred. No. 1.1e+03;
Matches 4; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 ALNVLN 6
Db |||||
72 ALNVLV 77

RESULT 36
Q98461 PRELIMINARY; PRT; 82 AA.
ID Q98461
AC Q98461;
DT 01-FEB-1997 (TrEMBLrel. 02, Created)
DT 01-FEB-1997 (TrEMBLrel. 02, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE A409R protein.
GN A409R.
OS Paramycium bursaria chlorella virus 1 (PBCV-1).
OC Viruses; dsDNA viruses, no RNA stage; Phycodnaviridae; Chlorovirus.
OX NCBI_TaxID=10506;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=96400190; PubMed=8806566;
RA Kutish G.F., Li Y., Lu Z., Furuta M., Rock D.L.; Van Etten J.L.;
RT "Analysis of 76 kb of the chlorella virus PBCV-1 330-kb genome: map
positions 182 to 258.";
RL Virology 223:303-317(1996).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=20013326; PubMed=10544099;
RA Kaiser A., Vollmert M., Tholl D., Furuta M., Rock D.L.; Van Etten J.L.;
RA Lisee A.D., Nickerson K.W., Van Etten J.L.;
RT "Chlorella virus PBCV-1 encodes a functional homospermidine
synthase.";
RL Virology 263:254-262(1999).
RN [3]
RP SEQUENCE FROM N.A.
RX MEDLINE=20478054; PubMed=11021991;
RA Sun L., Gurnon J.R., Adams B.J., Graves M.V., Van Etten J.L.;
RT "Characterization of a beta-1,3-glucanase encoded by chlorella virus
PBCV-1.";
RL Virology 276:27-36(2000).
RN [4]
RP SEQUENCE FROM N.A.
RA Van Etten J.L.;
RL Submitted (DEC-1995) to the EMBL/GenBank/DBJ databases.
RN [5]
RP SEQUENCE FROM N.A.
RA Van Etten J.L.;
RL Submitted (MAY-1997) to the EMBL/GenBank/DBJ databases.
RN [6]
RP SEQUENCE FROM N.A.
RA Van Etten J.L.;
RL Submitted (SEP-1997) to the EMBL/GenBank/DBJ databases.
RN [7]
RP SEQUENCE FROM N.A.
RA Graves M.V., Van Etten J.L.;
RL Submitted (DEC-1998) to the EMBL/GenBank/DBJ databases.
```



```

DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE Escherichia coli.
OS Hypothetical protein.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Escherichia.
OX NCBI_TaxID=562;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=493/89;
RA Janka A.;
RT "Sorbitol fermenting Shiga toxin producing Escherichia coli O157:H-
RT contains homologues of the Shigella resistance locus (SRL)
RT pathogenicity island-sequence of Shigella flexneri 2a.";
RL Submitted (DEC-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AJ534392; CAD58981.1; -.
KW Hypothetical protein.
SQ SEQUENCE 90 AA; 10424 MW; 6171FAB686992C32 CRC64;

Query Match 71.0%; Score 22; DB 2; Length 90;
Best Local Similarity 80.0%; Pred. No. 1.2e+03;
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 3 NVLMA 7
DB 20 NILMA 24

RESULT 41
Q8XYQ6 PRELIMINARY; PRT; 91 AA.
AC Q8XYQ6; 2002 (TrEMBLrel. 20, Created)
DT 01-MAR-2002 (TrEMBLrel. 20, Last sequence update)
DT 01-MAR-2002 (TrEMBLrel. 20, Last annotation update)
DE Probable transmembrane protein.
GN RSC1702 OR RS02891.
OS Ralstonia solanacearum (Pseudomonas solanacearum).
OC Bacteria; Proteobacteria; Betaproteobacteria; Burkholderiales;
OC Ralstoniaceae; Ralstonia.
OX NCBI_TaxID=305;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=GM11000;
RX MEDLINE=21681879; PubMed=11823852;
RA Salanoubat M., Genin S., Artiguenave F., Gouzy J., Mangenot S.,
RA Ariat M., Billault A., Brottier P., Camus J.C., Cattolico L.,
RA Chandler M., Choisme N., Claudel-Renard C., Cunnac S., Demange N.,
RA Gaopin C., Lavie M., Moisan A., Robert C., Saurin W., Schiex T.,
RA Sigquier P., Thebaud P., Whalen M., Wincker P., Levy M.,
RA Weissenbach J., Boucher C.A.;
RT "Genome sequence of the plant pathogen Ralstonia solanacearum.";
RL Nature 415:497-502(2002).
DR EMBL; AL646066; CAD15404.1; -.
KW Complete proteome.
SQ SEQUENCE 91 AA; 9778 MW; 431B829AC56608FA CRC64;

Query Match 71.0%; Score 22; DB 16; Length 91;
Best Local Similarity 71.4%; Pred. No. 1.2e+03;
Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 ALNVLMA 7
DB 61 ALNVAVA 67

RESULT 42
Q8NL30 PRELIMINARY; PRT; 91 AA.
AC Q8NL30;
DT 01-OCT-2002 (TrEMBLrel. 22, Created)
DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
DT 01-OCT-2002 (TrEMBLrel. 22, Last annotation update)

```

```

DE Phage-related protein.
GN ORF90 OR XAC2649 OR XCC2977.
OS Xanthomonas axonopodis (pv. citri), and
OS Xanthomonas campestris (pv. campestris).
OC Bacteria; Proteobacteria; Gammaproteobacteria; Xanthomonadales;
OC Xanthomonadaceae; Xanthomonas.
OX NCBI_TaxID=92829, 340;
RN [1]
RP SEQUENCE FROM N.A.
RC SPECIES=X.a.citri, and X.c.campestris;
RC STRAIN=306 / ATCC 13902 / XV 101, and ATCC 33913 / NCPPB 528;
RX MEDLINE=22022145; PubMed=12024217;
RA da Silva A.C.R., Ferro J.A., Reinach F.C., Farah C.S., Furlan L.R.,
RA Quaggio R.B., Monteiro-Vitorello C.B., Van Sluys M.A., Almeida N.F.,
RA Alves L.M.C., do Amaral A.M., Bertolini M.C., Camargo L.E.A.,
RA Camarotte G., Cannavan F., Cardozo J., Chambergro F., Ciapina L.P.,
RA Cicarelli R.M.B., Coutinho L.L., Cursino-Santos J.R., El-Dorry H.,
RA Faria J.B., Ferreira A.J.S., Ferreira R.C.C., Ferro M.I.T.,
RA Formighieri E.F., Franco M.C., Greggio C.C., Gruber A.,
RA Katayama A.M., Kishi L.T., Leite R.P., Lemos E.G.M., Lemos M.V.F.,
RA Locati E.C., Machado M.A., Madeira A.M.B.N., Martinez-Rossi N.M.,
RA Martins E.C., Meidanis J., Menck C.F.M., Miyaki C.Y., Moon D.H.,
RA Moreira L.M., Novo M.T.M., Okura V.K., Oliveira M.C., Oliveira V.R.,
RA Pereira H.A., Rossi A., Sena J.A.D., Silva C., de Souza R.F.,
RA Spinola L.A.F., Takita M.A., Tamura R.E., Teixeira E.C., Tezza R.I.D.,
RA Trindade dos Santos M., Truffi D., Tsai S.M., White F.F.,
RA Setubal J.C., Kitajima J.P.;
RT "Comparison of the genomes of two Xanthomonas pathogens with differing
RT host specificities.";
RL Nature 417:459-463(2002).
DR EMBL; AB011905; AAM37496.1; -.
DR EMBL; AB012412; AAM42249.1; -.
KW Complete proteome.
SQ SEQUENCE 91 AA; 9838 MW; 531AR81CA1F71F3B CRC64;

Query Match 71.0%; Score 22; DB 16; Length 91;
Best Local Similarity 57.1%; Pred. No. 1.2e+03;
Matches 4; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 1 ALNVLMA 7
DB 50 ALHILMA 56

RESULT 43
Q93F09 PRELIMINARY; PRT; 92 AA.
AC Q93F09;
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE Hypothetical 10.7 kDa protein.
OS Shigella flexneri 2a.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Shigella.
OX NCBI_TaxID=42897;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=21437601; PubMed=11553538;
RA Luck S.N., Turner S.A., Rajakumar K., Sakellaris H., Adler B.;
RT "Ferric Dicitrate Transport System (Fec) of Shigella flexneri 2a
RT YSH6000 Is Encoded on a Novel Pathogenicity Island Carrying Multiple
RT Antibiotic Resistance Genes.";
RL Infect. Immun. 69:6012-6021(2001).
DR EMBL; AF326777; AAL08466.1; -.
KW Hypothetical protein.
SQ SEQUENCE 92 AA; 10730 MW; 086451277558E0DD CRC64;

Query Match 71.0%; Score 22; DB 2; Length 92;
Best Local Similarity 80.0%; Pred. No. 1.2e+03;
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 3 NVLMA 7

```

```
Db          22 NILWA 26          29 ALNWIM 34

Search completed: February 17, 2004, 10:56:27
Job time : 12.9109 secs

RESULT 44
Q8FG80
ID   Q8FG80      PRELIMINARY;      PRT;      92 AA.
AC   Q8FG80;
DT   01-MAR-2003 (TrEMBLrel. 23, Created)
DT   01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT   01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE   Hypothetical protein.
GN   C2505.
OS   Escherichia coli O6.
OC   Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC   Enterobacteriaceae; Escherichia.
OX   NCBI_TaxID=217992;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=O6:H1 / CFT073 / ATCC 700928;
RX   MEDLINE=22388234; PubMed=12471157;
RA   Welch R.A., Burland V., Plunkett G. III, Redford P., Roesch P.,
RA   Rasko D., Buckles E.L., Liou S.-R., Boutin A., Hackett J., Stroud D.,
RA   Mayhew G.F., Rose D.J., Zhou S., Schwartz D.C., Perna N.T.,
RA   Mobley H.L.T., Donnenberg M.S., Blattner F.R.;
RT   "Extensive mosaic structure revealed by the complete genome sequence
RT   of uropathogenic Escherichia coli.";
RL   Proc. Natl. Acad. Sci. U.S.A. 99:17020-17024(2002).
DR   EMBL; AE016762; AAN80961.1; -.
KW   Hypothetical protein; Complete proteome.
SQ   SEQUENCE 92 AA; 10652 MW; 7E32132B717EA2BF CRC64;

Query Match          71.0%; Score 22; DB 16; Length 92;
Best Local Similarity 80.0%; Pred. No. 1.2e+03;
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 3 NVLWA 7
Db          22 NILWA 26

RESULT 45
Q92JC7
ID   Q92JC7      PRELIMINARY;      PRT;      95 AA.
AC   Q92JC7;
DT   01-DEC-2001 (TrEMBLrel. 19, Created)
DT   01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT   01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE   Hypothetical protein RC0140.
GN   RC0140.
OS   Rickettsia conorii.
OC   Bacteria; Proteobacteria; Alphaproteobacteria; Rickettsiales;
OC   Rickettsiaceae; Rickettsiae; Rickettsia.
OX   NCBI_TaxID=781;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=Malish 7;
RX   MEDLINE=21442074; PubMed=11557893;
RA   Ogata H., Audic S., Renesto-Audiffren P., Fournier P.-E., Barbe V.,
RA   Samson D., Roux V., Cossart P., Weissenbach J., Claverie J.-M.,
RA   Raoult D.;
RT   "Mechanisms of evolution in Rickettsia conorii and R. prowazekii.";
RL   Science 293:2093-2098(2001).
DR   EMBL; AE008580; AAL02678.1; -.
KW   Hypothetical protein; Complete proteome.
SQ   SEQUENCE 95 AA; 10750 MW; 7253E0673CBA777C CRC64;

Query Match          71.0%; Score 22; DB 16; Length 95;
Best Local Similarity 66.7%; Pred. No. 1.3e+03;
Matches 4; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 ALNWLM 6
Db          22 NILWA 26
```

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: February 17, 2004, 10:50:13 ; Search time 2.21782 Seconds  
(without alignments)  
148.428 Million cell updates/sec

Title: US-09-900-147-7  
Perfect score: 31  
Sequence: 1 ALNVLMA 7

Scoring table:  
Gapop 10.0 , Gapext 0.5

Searched: 127863 seqs, 47026705 residues

Total number of hits satisfying chosen parameters: 13973

Minimum DB seq length: 0  
Maximum DB seq length: 100

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : SwissProt\_41.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	27	87.1	56	SCP2_MESMA	Q9njp7 mesobuthus
2	27	87.1	56	SCP3_MESMA	Q9u8dl mesobuthus
3	25	80.6	98	1 NULM_BALMU	P41301 balaenopter
4	24	77.4	42	1 RETS_BOVIN	P82708 bos taurus
5	24	77.4	94	1 YALIE_ECOLI	P36768 escherichia
6	24	77.4	98	1 NULM_BALPH	P24976 balaenopter
7	24	77.4	98	1 NULM_DASNO	O21333 dasypus nov
8	23	74.2	96	1 C553_HELPJ	Q9zjz9 helicobacte
9	23	74.2	96	1 YFRC_PROVU	P20927 proteus vul
10	22	71.0	87	1 YWCE_BAGSU	P39603 bacillus su
11	22	71.0	98	1 NULM_BOVIN	P03902 bos taurus
12	22	71.0	98	1 NULM_HIPAM	P26632 sus scrofa
13	22	71.0	98	1 NULM_PIG	O8754 ovis aries
14	22	71.0	98	1 NULM_SHEEP	Q22702 caenorhabdi
15	21	67.7	57	1 YCUE_CABEL	Q22702 caenorhabdi
16	21	67.7	63	1 Y697_ARCFU	Q15513 homo sapien
17	21	67.7	63	1 SPHA_HUMAN	Q9v0v8 pyrococcus
18	21	67.7	75	1 RUXH_PYRAB	O74016 pyrococcus
19	21	67.7	75	1 RUXX_PYRHO	Q9u0p4 pyrococcus
20	21	67.7	76	1 RUXX_PYRHO	Q9u0p4 pyrococcus
21	21	67.7	89	1 EFIB_METH	P22889 octodon deg
22	21	67.7	91	1 TAPP_OCTDE	P25757 serratia ma
23	21	67.7	91	1 YOHL_SERMA	Q8w9m8 dugong dugo
24	21	67.7	98	1 NULM_DUGDU	Q36457 ornithorhyn
25	21	67.7	98	1 NULM_ORNAN	P23320 synchococc
26	20	64.5	29	1 PSAX_SYNVU	Q9ppp4 ureaplasma
27	20	64.5	77	1 ACPH_UREFA	P49394 borrelia bu
28	20	64.5	85	1 RS20_BORBU	P00114 synchococc
29	20	64.5	87	1 CYC6_SYNLI	P00115 synchococc
30	20	64.5	87	1 CYC6_SYNP3	Q9myw0 bos taurus
31	20	64.5	89	1 TRP6_BOVIN	Q18060 caenorhabdi
32	20	64.5	91	1 ILGI_CAEEL	O67242 aquifex ae
33	20	64.5	93	1 FLIE_AQUAE	

34	20	64.5	94	1 YGBE_ECOLI	P45956 escherichia
35	20	64.5	95	1 YD82_ARCFU	O28889 archaeoglob
36	20	64.5	98	1 NULM_DIDMA	P41307 didelphis m
37	20	64.5	98	1 NULM_MACRO	P92667 macropus ro
38	19	61.3	25	1 IPYR_PSEAN	P80898 pseudanabae
39	19	61.3	33	1 PK1_DICDI	P34101 dictyosteli
40	19	61.3	43	1 PSAX_ANASP	P58566 anabaena sp
41	19	61.3	52	1 Y567_PSEAE	Q915w9 pseudomonas
42	19	61.3	85	1 COXE_BOVIN	P13182 bos taurus
43	19	61.3	86	1 C555_CHLLT	P00123 chlorobium
44	19	61.3	86	1 YM17_MARPO	P38460 marchantia
45	19	61.3	87	1 SYNC_SACDO	P41908 saccharomyc

## ALIGNMENTS

## RESULT 1

SCP2\_MESMA STANDARD; PRT; 56 AA.

AC Q9NJ7; P58491;  
DT 28-FEB-2003 (Rel. 41, Created)  
DT 28-FEB-2003 (Rel. 41, Last sequence update)  
DT 15-SEP-2003 (Rel. 42, Last annotation update)  
DE Neurotoxin Bmp02 precursor (Potassium ion channel blocker P02P).  
OS Mesobuthus martensii (Manchurian scorpion) (Buthus martensii).  
OC Eukaryota; Metazoa; Arthropoda; Chelicerata; Arachnida; Scorpiones;  
OC Buthidae; Mesobuthus.  
OX NCBI\_TaxID=34649;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC TISSUE=Venom gland;  
RX MEDLINE=99402983; PubMed=10471839;  
RA Zhu S.-Y., Li W.-X., Zeng X.-C., Jiang D.-H., Mao X., Liu H.;  
RT "Molecular cloning and sequencing of two 'short chain' and two 'long chain' K(+) channel-blocking peptides from the Chinese scorpion Buthus martensii Karsch.";  
RT FEBS Lett. 457:509-514 (1999).  
RN [2]  
RP STRUCTURE BY NMR OF 29-56.  
RC TISSUE=Venom;  
RX MEDLINE=20530297; PubMed=11076505;  
RA Xu Y.O., Wu J.H., Pei J.M., Shi Y.Y., Ji Y.H., Tong Q.C.;  
RT "Solution structure of Bmp02, a new potassium channel blocker from the venom of the Chinese scorpion Buthus martensii Karsch.";  
RT Biochemistry 39:13669-13675 (2000).  
CC -1- FUNCTION: Blocks potassium channels.  
CC -1- SUBCELLULAR LOCATION: Secreted.  
CC -1- TISSUE SPECIFICITY: Expressed by the venom gland.  
CC -1- SIMILARITY: BELONGS TO THE SCORPION LEIUTOXIN FAMILY.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----

DR EMBL; AF132975; AAF31296.1; -  
DR PDB; 1DU9; 28-MAR-01.  
KW Toxin; Neurotoxin; Ionic channel inhibitor;  
KW Potassium channel inhibitor; Signal; 3D-structure.  
FT SIGNAL 1 28  
FT CHAIN 29 56 NEUROTOXIN Bmp02.  
FT DISULFID 31 47  
FT DISULFID 34 52  
FT DISULFID 38 54  
SQ SEQUENCE 56 AA; 6015 MW; 70953032042F8672 CRC64;  
-----

Query Match 87.1%; Score 27; DB 1; Length 56;  
Best Local Similarity 71.4%; Pred. No. 5.5;  
Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Tue Feb 17 11:55:57 2004

us-09-900-147-7-rsp

```

QY      1 ALNVLM 7
Db      13 AMNVMA 19

RESULT 2
SCP3 MESMA STANDARD; PRT; 56 AA.
ID SCP3 MESMA STANDARD; PRT; 56 AA.
AC QUBDI; 2003 (Rel. 41, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DE Neurotoxin Bmp03 precursor (Potassium ion channel blocker P03).
OS Mesobuthus martensii (Mandchurian scorpion) (Buthus martensii).
OC Eukaryota; Metazoa; Arthropoda; Chelicerata; Arachnida; Scorpiones;
OC Buthoidea; Buthidae; Mesobuthus.
OX NCBI_TaxID=34649;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Venom gland;
RX MEDLINE=99311393; PubMed=10386622;
RA Wu J.J.; Dai L.; Lan Z.D.; Chi C.-W.;
RT "Genomic organization of three neurotoxins active on small conductance
RT Ca2+-activated potassium channels from the scorpion Buthus martensi
RT Karsch.";
RL FEBS Lett. 452:360-364 (1999).
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=Venom gland;
RX MEDLINE=99402983; PubMed=10471839;
RA Zhu S.-Y.; Li W.-X.; Zeng X.-C.; Jiang D.-H.; Mao X.; Liu H.;
RT "Molecular cloning and sequencing of two 'short chain' and two 'long
RT chain' K(+) channel-blocking peptides from the Chinese scorpion Buthus
RT martensii Karsch.";
RL FEBS Lett. 457:509-514 (1999).
CC -!- FUNCTION: Blocks small conductance Calcium-activated potassium
CC channels.
CC -!- SUBCELLULAR LOCATION: Secreted.
CC -!- TISSUE SPECIFICITY: Expressed by the venom gland.
CC -!- SIMILARITY: BELONGS TO THE SCORPION LEIUTOXIN FAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; AF097408; AAF01253.1; -.
CC DR EMBL; AF156170; AAF29463.1; -.
CC DR HSSP; Q9NJP7; 1DU9.
CC KW Toxin; Neurotoxin; Ionic channel inhibitor;
CC KW Potassium channel inhibitor; Signal.
FT SIGNAL 1 28 BY SIMILARITY.
FT CHAIN 29 56 NEUROTOXIN BMP03.
FT DISULFID 31 47 BY SIMILARITY.
FT DISULFID 34 52 BY SIMILARITY.
FT DISULFID 38 54 BY SIMILARITY.
SQ SEQUENCE 56 AA; 6001 MW; 70953032007E8672 CRC64;

Query Match 87.1%; Score 27; DB 1; Length 56;
Best Local Similarity 71.4%; Pred. No. 5.5;
Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY      1 ALNVLM 7
Db      13 AMNVMA 19

RESULT 3
NULM_BALMU

Query Match 80.6%; Score 25; DB 1; Length 98;
Best Local Similarity 83.3%; Pred. No. 31;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY      2 LNVLM 7
Db      6 MNVLM 11

RESULT 4
RETS BOVIN STANDARD; PRT; 42 AA.
ID RETS BOVIN STANDARD; PRT; 42 AA.
AC P82708;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Retinol-binding protein III, cellular (CRBP-III) (Fragment).
GN RBP5.
OS Bos taurus (Bovine).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Euthera; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
OC Bovidae; Bovinae; Bos.
OX NCBI_TaxID=9913;
RN [1]
RP SEQUENCE, AND FUNCTION.
RC TISSUE=Kidney;
RX MEDLINE=21173623; PubMed=11274389;
RA Polli C.; Calderone V.; Ottonello S.; Bolchi A.; Zanotti G.;
RA Stoppini M.; Berni R.;
RT "Identification, retinoid binding and X-ray analysis of a human
RT retinol-binding protein.";
RL Proc. Natl. Acad. Sci. U.S.A. 98:3710-3715 (2001).
CC -!- FUNCTION: Intracellular transport of retinol.
CC -!- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).

```



```

CC -1- TISSUE SPECIFICITY: Kidney.
CC -1- SIMILARITY: BELONGS TO THE FABP/P2/CRBP/CRABP FAMILY OF
CC TRANSPORTERS.
DR HSSP; P82980; 1GGL.
DR InterPro; IPR000463; Fatty acid BP.
DR InterPro; IPR000566; Lipocalin_cyFABP.
DR Pfam; PF00061; lipocalin; 1.
DR PROSITE; PS00214; FABP; FALSE NEG.
KW Vitamin A; Retinol-binding; Transport.
FT NON TER 42
SQ SEQUENCE 42 AA; 4892 MW; ACB4F1399FDD7F09 CRC64;

Query Match 77.4%; Score 24; DB 1; Length 42;
Best Local Similarity 85.7%; Pred. No. 23;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 ALNVLMA 7
Db 22 ALNVNMA 28
|||||

RESULT 5
YATE ECOLI
ID YATE ECOLI STANDARD; PRT; 94 AA.
AC P36768; P77343;
DT 01-JUN-1994 (Rel. 29, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Hypothetical protein yaeE
GN YAE OR B0391 OR C0499 OR Z0487 OR ECS0441 OR SF0327.
OS Escherichia coli,
OS Escherichia coli O6,
OS Escherichia coli O157:H7, and
OS Shigella flexneri.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Escherichia.
OX NCBI_TaxID=562, 217992, 83334, 623;
RN [1]
RP SEQUENCE FROM N.A.
RC SPECIES=E.coli; STRAIN=K12;
RX MEDLINE=96400909; PubMed=8807285;
RA Ryder L., Sharples G.J., Lloyd R.G.;
RT "Recombination-dependent growth in exonuclease-depleted recBC sbcBC
strains of Escherichia coli K-12.";
RL Genetics 143:1101-1114(1996).
RN [2]
RP SEQUENCE FROM N.A.
RC SPECIES=E.coli; STRAIN=K12 / MG1655;
RX MEDLINE=97426617; PubMed=9278503;
RA Blattner F.R., Plunkett G. III, Bloch C.A., Perna N.T., Burland V.,
Riley M., Collado-Vides J., Glasner J.D., Rode C.K., Mayhew G.F.,
Gregor J., Davis N.W., Kirkpatrick H.A., Goeden M.A., Rose D.J.,
Mau B., Shao Y.;
RT "The complete genome sequence of Escherichia coli K-12.";
RL Science 277:1453-1474(1997).
RN [3]
RP SEQUENCE FROM N.A.
RC SPECIES=E.coli;
RA Duncan M., Allen E., Araujo R., Aparicio A.M., Chung E., Davis K.,
Federle N., Hyman R., Kalman S., Komp C., Kurdi O., Lew H., Lin D.,
Namath A., Oefner P., Roberts D., Schramm S., Davis R.W.;
RL Submitted (NOV-1996) to the EMBL/GenBank/DBJ databases.
RN [4]
RP SEQUENCE FROM N.A.
RC SPECIES=E.coli; STRAIN=O6:H1 / CFT073 / ATCC 700928;
RX MEDLINE=22388234; PubMed=12471157;
RA Welch R.A., Burland V., Plunkett G. III, Redford P., Roesch P.,
Rasko D., Buckles E.L., Liou S.-R., Boutin A., Hackett J., Stroud D.,
Mayhew G.F., Rose D.J., Zhou S., Schwartz D.C., Perna N.T.,
Mabley H.L.T., Donnenberg M.S., Blattner F.R.;
RT "Extensive mosaic structure revealed by the complete genome sequence
of uropathogenic Escherichia coli.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:17020-17024(2002).

[5]
RN RP SEQUENCE FROM N.A.
RC SPECIES=E.coli; STRAIN=O157:H7 / EDL933 / ATCC 700927;
RX MEDLINE=21074935; PubMed=11206551;
RA Perna N.T., Plunkett G. III, Burland V., Mau B., Glasner J.D.,
Rose D.J., Mayhew G.F., Evans P.S., Gregor J., Kirkpatrick H.A.,
Posfai G., Hackett J., Klink S., Boutin A., Shao Y., Miller L.,
Grotbeck E.J., Davis N.W., Lim A., Dimalanta E.T., Potamousis K.,
Apodaca J., Anantharaman T.S., Lin J., Yen G., Schwartz D.C.,
Welch R.A., Blattner F.R.;
RT "Genome sequence of enterohaemorrhagic Escherichia coli O157:H7.";
RL Nature 409:529-533(2001).
RN [6]
RN RP SEQUENCE FROM N.A.
RC SPECIES=E.coli; STRAIN=O157:H7 / RIMD 0509952;
RX MEDLINE=21156231; PubMed=11258796;
RA Havaashi T., Makino K., Ohnishi M., Kurokawa K., Ishii K., Yokoyama K.,
Hara C.-G., Ohtsubo E., Nakayama K., Murata T., Tanaka M., Tobe T.,
Iida T., Takami H., Honda T., Sasakawa C., Ogasawara N., Yasunaga T.,
Kuhara S., Shiba T., Hattori M., Shinagawa H.;
RT "Complete genome sequence of enterohemorrhagic Escherichia coli
O157:H7 and genomic comparison with a laboratory strain K-12.";
RL DNA Res. 8:11-22(2001).
RN [7]
RN RP SEQUENCE FROM N.A.
RC SPECIES=S.flexneri; STRAIN=301 / Serotype 2a;
RX MEDLINE=22272406; PubMed=12384590;
RA Jin O., Yuan Z., Xu J., Wang Y., Shen Y., Lu W., Wang J., Liu H.,
Yang J., Yang F., Zhang X., Zhang J., Yang G., Wu H., Qu D., Dong J.,
Sun L., Xue Y., Zhao A., Gao Y., Zhu J., Kan B., Ding K., Chen S.,
Cheng H., Yao Z., He B., Chen R., Ma D., Qiang B., Wen Y., Hou Y.,
Yu J.;
RT "Genome sequence of Shigella flexneri 2a: insights into pathogenicity
through comparison with genomes of Escherichia coli K12 and O157.";
RL Nucleic Acids Res. 30:4432-4441(2002).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).
CC -----
DR EMBL; X76979; CAA54286.1; -
DR EMBL; AE000145; AAC73494.1; -
DR EMBL; U73857; AAB18115.1; -
DR EMBL; AE016756; AAN78977.1; -
DR EMBL; AE005218; AAG54737.1; -
DR EMBL; AP002551; BAB33864.1; -
DR EMBL; AE015066; AAN41986.1; -
DR PIR; A90684; A30684.
DR PIR; E85534; E85534.
DR PIR; G64767; G64767.
DR EcoGene; EG12159; yaeE.
KW Hypothetical protein; Complete proteome.
FT CONFLICT 77 94 SEFHQVAEPTSLCRYL -> TVSFICKLPNPPLCAAI
FT NSPSP (IN REF. 1).
SQ SEQUENCE 94 AA; 10234 MW; D7EF5COAFD86D661 CRC64;

Query Match 77.4%; Score 24; DB 1; Length 94;
Best Local Similarity 83.3%; Pred. No. 53;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 ALNVLM 6
Db 51 ALNVLL 56
|||||

RESULT 6
NULM BALPH STANDARD; PRT; 98 AA.
ID NULM BALPH
AC P24976;

```

use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

CC 01-MAR-1992 (Rel. 21, Created)  
CC 01-MAR-1992 (Rel. 21, Last sequence update)  
CC 16-OCT-2001 (Rel. 40, Last annotation update)  
CC NADH-ubiquinone oxidoreductase chain 4L (EC 1.6.5.3).  
CC MTND4L OR ND4L OR NADH4L.  
CC Balaenoptera physalus (Finback whale) (Common rorqual).  
CC Mitochondrion.  
CC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
CC Mammalia; Eutheria; Cetartiodactyla; Cetacea; Mysticeti;  
CC Balaenopteridae; Balaenoptera.  
CC NCBI\_TaxID=9770;  
CC [1]  
CC SEQUENCE FROM N.A.  
CC STRAIN=Isolate No. 27 / Anno 1987; TISSUE=Liver;  
CC MEDLINE=92139449; PubMed=1779436;  
CC Arnason U., Gullberg A., Widgren B.;  
CC "The complete nucleotide sequence of the mitochondrial DNA of the fin  
CC whale, Balaenoptera physalus."  
CC J. Mol. Evol. 33:556-568(1991).  
CC -1- CATALYTIC ACTIVITY: NADH + ubiquinone = NAD(+) + ubiquinol.  
CC  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

CC EMBL; X61145; CAA43447.1; -  
CC PIR; A58851; S24920.  
CC InterPro; IPR001133; Oxidored\_4L.  
CC InterPro; IPR003214; Oxidred4L.  
CC Pfam; PF00420; oxidored\_g2; 1.  
CC Pfam; PD000359; Oxidred4L; 1.  
CC Oxidoreductase; NAD; Ubiquinone; Mitochondrion.  
CC SEQUENCE 98 AA; 10763 MW; 0070D3D6C3AE805F CRC64;  
CC  
CC Query Match 77.4%; Score 24; DB 1; Length 98;  
CC Best Local Similarity 66.7%; Pred. No. 55;  
CC Matches 4; Conservative 2; Mismatches 0; Indels 0; Gaps 0;  
CC  
CC QY 2 LNVLMA 7  
CC DB 6 MNILMA 11  
CC  
CC RESULT 7  
CC ID NULM\_DASNO STANDARD; PRT; 98 AA.  
CC AC 021333;  
CC DT 15-JUL-1998 (Rel. 36, Created)  
CC DT 15-JUL-1998 (Rel. 36, Last sequence update)  
CC DT 15-JUL-1998 (Rel. 36, Last annotation update)  
CC DE NADH-ubiquinone oxidoreductase chain 4L (EC 1.6.5.3).  
CC GN MTND4L OR ND4L OR NADH4L.  
CC OS Dasypus novemcinctus (Nine-banded armadillo).  
CC Mitochondrion.  
CC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
CC Mammalia; Eutheria; Edentata; Dasypodidae; Dasypus.  
CC NCBI\_TaxID=9361;  
CC [1]  
CC SEQUENCE FROM N.A.  
CC MEDLINE=97357423; PubMed=9214749;  
CC Arnason U., Gullberg A., Janke A.;  
CC "Phylogenetic analyses of mitochondrial DNA suggest a sister group  
CC relationship between Xenarthra (Edentata) and Ferungulates."  
CC Mol. Biol. Evol. 14:762-768(1997).  
CC -1- CATALYTIC ACTIVITY: NADH + ubiquinone = NAD(+) + ubiquinol.  
CC  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its

use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

CC EMBL; Y11832; CAA72524.1; -  
CC PIR; T11449; T11449.  
CC InterPro; IPR001133; Oxidored\_4L.  
CC InterPro; IPR003214; Oxidred4L.  
CC Pfam; PF00420; oxidored\_g2; 1.  
CC Pfam; PD000359; Oxidred4L; 1.  
CC Oxidoreductase; NAD; Ubiquinone; Mitochondrion.  
CC SEQUENCE 98 AA; 10840 MW; D0FF9BC309048774 CRC64;  
CC  
CC Query Match 77.4%; Score 24; DB 1; Length 98;  
CC Best Local Similarity 66.7%; Pred. No. 55;  
CC Matches 4; Conservative 2; Mismatches 0; Indels 0; Gaps 0;  
CC  
CC QY 2 LNVLMA 7  
CC DB 6 LNILMA 11  
CC  
CC RESULT 8  
CC ID\_C553\_HELPJ STANDARD; PRT; 96 AA.  
CC AC 092J29;  
CC DT 28-FEB-2003 (Rel. 41, Created)  
CC DT 28-FEB-2003 (Rel. 41, Last sequence update)  
CC DT 28-FEB-2003 (Rel. 41, Last annotation update)  
CC DE Cytochrome c-553 precursor (C553).  
CC GN JHP1148.  
CC OS Helicobacter pylori J99 (Campylobacter pylori J99).  
CC Bacteria; Proteobacteria; Epsilonproteobacteria; Campylobacteriales;  
CC Helicobacteraceae; Helicobacter.  
CC NCBI\_TaxID=85963;  
CC RN [1]  
CC SEQUENCE FROM N.A.  
CC MEDLINE=99120557; PubMed=9923682;  
CC Alm R.A., Ling L.-S.L., Moir D.T., King B.L., Brown E.D., Doig P.C.,  
CC Smith D.R., Noonan B., Guild B.C., deJonge B.L., Carmel G.,  
CC Tummino P.J., Caruso A., Uria-Nickelsen M., Mills D.M., Ives C.,  
CC Gibson R., Merberg D., Mills S.D., Jiang Q., Taylor D.E., Vovis G.F.,  
CC Trust T.J.;  
CC "Genomic sequence comparison of two unrelated isolates of the human  
CC gastric pathogen Helicobacter pylori."  
CC Nature 397:176-180(1999).  
CC -1- FUNCTION: Natural electron acceptor for a formate dehydrogenase.  
CC -1- PFM: BINDS ONE HEME GROUP PER MOLECULE.  
CC -1- SIMILARITY: BELONGS TO THE CLASS I CYTOCHROME C FAMILY.  
CC  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

CC EMBL; AE001542; AAD06721.1; -  
CC PIR; F71843; F71843.  
CC HSP; P04032; 2DVH.  
CC InterPro; IPR003088; Cyt\_C1.  
CC InterPro; IPR002329; Cyt\_C1C.  
CC InterPro; IPR000345; CytC\_heme\_bind.  
CC Pfam; PF00034; cytochrome\_c; 1.  
CC PRINTS; PR00605; CYTOCHROME\_C.  
CC PRODOM; PD00402; Cyt\_C\_bact; 1.  
CC PROSITE; PS00190; CYTOCHROME\_C; 1.  
CC Electron transport; Heme; Signal; Complete proteome.  
CC SIGNAL 19  
CC CHAIN 20 96  
CC CYTOCHROME C-553.

```

FT BINDING 29 29 HEME (COVALENT).
FT BINDING 32 32 HEME (COVALENT).
FT METAL 33 33 IRON (HEME AXIAL LIGAND) (BY
FT METAL 33 33 SIMILARITY).
FT METAL 73 73 IRON (HEME AXIAL LIGAND) (BY
FT METAL 73 73 SIMILARITY).
SQ SEQUENCE 96 AA; 10354 MW; 3E607AE5D422AD82 CRC64;

Query Match
Best Local Similarity 74.2%; Score 23; DB 1; Length 96;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 NVLMA 7
DB 15 NVLMA 19

RESULT 9
YFRC_PROVU STANDARD; PRT; 96 AA.
AC P20927;
DT 01-FEB-1991 (Rel. 17, Created)
DT 01-FEB-1991 (Rel. 17, Last sequence update)
DE FD operon hypothetical protein C.
OS Proteus vulgaris.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Proteus.
OX NCBI_TaxID=585;
RN [1]
SEQUENCE FROM N.A.
RX MEDLINE=88004470; PubMed=3308458;
RA Cole S.T.;
RT "Nucleotide sequence and comparative analysis of the frd operon
RT encoding the fumarate reductase of Proteus vulgaris. Extensive
RT sequence divergence of the membrane anchors and absence of an
RT frd-linked ampC cephalosporinase gene.";
RL Eur. J. Biochem. 167:481-488(1987).
CC -1- SIMILARITY: BELONGS TO THE HUPF/HYPC FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; X06151; CAA29511.1; -.
CC PIR; S00119; S00119.
CC InterPro; IPR001109; HupF_HypC.
CC Pfam; PF01455; HupF_HypC; 1.
CC PIRSF; PIRSF005618; HupF_HypC; 1.
CC PRINTS; PR00445; HUPFHYP.
CC ProDom; PD003112; HupF_HypC; 1.
CC TIGRFAMs; TIGR00074; hupC_hupF; 1.
CC PROSITE; PS01097; HUPF_HYP; 1.
CC Hypothetical protein.
SQ SEQUENCE 96 AA; 10300 MW; EC094FLF37956EB3 CRC64;

Query Match
Best Local Similarity 74.2%; Score 23; DB 1; Length 96;
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 LNVLMA 7
DB 69 LNALMA 74

RESULT 10
YWCE_BAGSU STANDARD; PRT; 87 AA.
ID YWCE_BAGSU
AC P39603;

```

```

DT 01-FEB-1995 (Rel. 31, Created)
DT 01-FEB-1995 (Rel. 31, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Hypothetical protein ywce precursor.
GN YWCE OR IPA-41R.
OS Bacillus subtilis.
OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.
OX NCBI_TaxID=1423;
RN [1]
SEQUENCE FROM N.A.
RP STRAIN=168;
RX MEDLINE=95020537; PubMed=7934828;
RA Glaser P., Kunst F., Arnaud M., Coudart M.P., Gonzales W.,
RA Hullo M.F., Ionescu M., Lubochinsky B., Marcelino L., Moszer I.,
RA Presecan E., Santana M., Schneider E., Schweizer J., Vertes A.,
RA Rapoport G., Danchin A.;
RT "Bacillus subtilis genome project: cloning and sequencing of the 97
RT kb region from 325 degrees to 333 degrees.";
RL Mol. Microbiol. 10:371-384(1993).
RN [2]
SEQUENCE FROM N.A.
RP STRAIN=168;
RX MEDLINE=98044033; PubMed=9384377;
RA Kunst F., Ogasawara N., Moszer I., Albertini A.M., Alloni G., S.,
RA Azevedo V., Bertero M.G., Bessieres P., Bolotin A., Borchert S.,
RA Borris R., Boursier L., Brans A., Braun M., Brignell S.C., Bron S.,
RA Broutillet S., Bruschi C.V., Caldwell B., Capuano V., Carter N.M.,
RA Choi S.K., Codani J.J., Connerton I.F., Cummings N.J., Daniel R.A.,
RA Denizot F., Devine K.M., Dusterhoft A., Ehrlich S.D., Emerson P.T.,
RA Entian K.D., Errington J., Fabret C., Fehrlich E., Foulger D.,
RA Fritz C., Fujita M., Fujita Y., Fuma S., Gallizzi A., Galleron N.,
RA Ghim S.Y., Glaser P., Goffeau A., Golightly E.J., Grandi G.,
RA Giuseppe G., Guy B.J., Haga K., Hatesch J., Harwood C.R., Henaut A.,
RA Hilbert H., Holsappel S., Hosono S., Hullo M.F., Itaya M., Jones L.,
RA Joris B., Karamata D., Kasahara Y., Klaerr-Blanchard M., Klein C.,
RA Kobayashi Y., Koetter P., Koningstein G., Krogh S., Kumano M.,
RA Kurita K., Lapidus A., Lardinois S., Lauber J., Lazarevic V.,
RA Lee S.M., Levine A., Liu H., Masuda S., Mauel C., Medigue C.,
RA Medina N., Mellado R.P., Mizuno M., Moestl D., Nakai S., Noback M.,
RA Noone D., O'Reilly M., Ogawa K., Ogiwara A., Oudega B., Park S.H.,
RA Parro V., Pohl T.M., Portetelle D., Porwollik S., Prescott A.M.,
RA Presecan E., Pujic P., Purnelle B., Rapoport G., Rey M., Reynolds S.,
RA Rieger M., Rivolta C., Rocha E., Roche B., Rose M., Sadaie Y.,
RA Sato T., Scanlan E., Schleich S., Schroeter R., Scoffone F.,
RA Sekiguchi J., Sekowska A., Seror S.J., Seror P., Shin B.S., Soldo B.,
RA Sorokin A., Tacconi E., Takagi T., Takahashi H., Takemaru K.,
RA Takeuchi M., Tanakoshi A., Tanaka T., Terpstra P., Tognoni A.,
RA Tosato V., Uchiyama S., Vandenbol M., Vannier F., Vaesarotti A.,
RA Viari A., Wambutt R., Wedler E., Wedler H., Weitzenecker T.,
RA Winters P., Wipat A., Yamamoto H., Yamane K., Yasumoto K., Yata K.,
RA Yoshida K., Yoshikawa H.F., Zumstein E., Yoshikawa H., Danchin A.;
RT "The complete genome sequence of the Gram-positive bacterium Bacillus
RT subtilis.";
RL Nature 390:249-256(1997).
CC -1- SUBCELLULAR LOCATION: Integral membrane protein (Potential).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; X73124; CAA51597.1; -.
CC EMBL; Z99123; CAB15839.1; -.
CC PIR; S39696; S39696.
CC Subtilist; BG10587; ywce.
CC Hypothetical protein; Transmembrane; Signal; Complete proteome.
FT SIGNAL 1 29 POTENTIAL.
FT CHAIN 30 87 HYPOTHETICAL PROTEIN YWCE.
FT TRANSMEM 30 46 POTENTIAL.
FT TRANSMEM 55 78 POTENTIAL.

```

```

SQ SEQUENCE 87 AA; 9959 MW; CC74135F423940CB CRC64;

Query Match 71.0%; Score 22; DB 1; Length 87;
Best Local Similarity 57.1%; Pred. No. 1.5e+02;
Matches 4; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 1 ALNVLMA 7
   |:|:|:|
Db 62 AVNVIVA 68

RESULT 11
NULM_BOVIN STANDARD; PRT; 98 AA.
AC P03902; Q8SFX7;
DT 21-JUL-1986 (Rel. 01, Created)
DT 21-JUL-1986 (Rel. 01, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE NADH-ubiquinone oxidoreductase chain 4L (EC 1.6.5.3).
GN MTND4L OR ND4L.
OS Bos taurus (Bovine).
OG Mitochondrion.
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovoidea;
OC Bovidae; Bovinae; Bos.
OX NCBI_TaxID=9913;
RN [1]
RP SEQUENCE FROM N.A.
RC MEDLINE=8310260; PubMed=7120390;
RA Anderson S., de Bruijn M.H.L., Coulson A.R., Eperon I.C., Sanger F.,
RA Young I.G.;
RT "Complete sequence of bovine mitochondrial DNA. Conserved features of
RT the mammalian mitochondrial genome.";
RL J. Mol. Biol. 156:683-717(1982).
[2]
RP SEQUENCE FROM N.A.
RC STRAIN=65, 66, D, and F;
RA Wetstein P.J.;
RA "Bos taurus mitochondrial protein coding regions.";
RL Submitted (MAR-2002) to the EMBL/GenBank/DBJ databases.
CC -1- CATALYTIC ACTIVITY: NADH + ubiquinone = NAD(+) + ubiquinol.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; V00654; CA24004.1; -
DR EMBL; J01394; AAB59276.1; -
DR EMBL; AF490528; AAM08325.1; -
DR EMBL; AF490529; AAM08338.1; -
DR EMBL; AF493541; AAM12797.1; -
DR EMBL; AF493542; AAM12810.1; -
DR PIR; A00429; OXB04L.
DR InterPro; IPR001133; Oxidored4L.
DR InterPro; IPR003214; Oxidred4L.
DR Pfam; PF00420; Oxidored_q2; 1.
DR ProDom; PD000359; Oxidred4L; 1.
DR Oxidoreductase; NAD; Ubiquinone; Mitochondrion.
KW VARIANT 65 65 V -> A (IN STRAIN 66).
SQ SEQUENCE 98 AA; 10797 MW; 4DB48B7DAS9C1881 CRC64;

Query Match 71.0%; Score 22; DB 1; Length 98;
Best Local Similarity 50.0%; Pred. No. 1.7e+02;
Matches 3; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 2 LNVLMA 7
   |:|:|:|
Db 6 MNILMA 11

RESULT 12
NULM_HIPAM STANDARD; PRT; 98 AA.
AC Q9ZZY3;
DT 30-MAY-2000 (Rel. 39, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE NADH-ubiquinone oxidoreductase chain 4L (EC 1.6.5.3).
GN MTND4L OR ND4L OR NADH4L.
OS Hippopotamus amphibius (Hippopotamus).
OG Mitochondrion.
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Hippopotamidae; Hippopotamus.
OX NCBI_TaxID=9833;
RN [1]
RP SEQUENCE FROM N.A.
RC MEDLINE=99098146; PubMed=9881471;
RA Ursing B.M., Arnason U.;
RA "Analyses of mitochondrial genomes strongly support a hippopotamus-
RA whale clade.";
RL Proc. R. Soc. Lond., B, Biol. Sci. 265:2251-2255(1998).
CC -1- CATALYTIC ACTIVITY: NADH + ubiquinone = NAD(+) + ubiquinol.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; AJ010957; CAA09436.1; -
DR InterPro; IPR001133; Oxidored4L.
DR InterPro; IPR003214; Oxidred4L.
DR Pfam; PF00420; Oxidored_q2; 1.
DR ProDom; PD000359; Oxidred4L; 1.
DR Oxidoreductase; NAD; Ubiquinone; Mitochondrion.
KW SEQUENCE 98 AA; 10776 MW; 5F095988CE0622F0 CRC64;

Query Match 71.0%; Score 22; DB 1; Length 98;
Best Local Similarity 50.0%; Pred. No. 1.7e+02;
Matches 3; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 2 LNVLMA 7
   |:|:|:|
Db 6 MNILMA 11

RESULT 13
NULM_PIG STANDARD; PRT; 98 AA.
AC P56632; 1998 (Rel. 37, Created)
DT 15-DEC-1998 (Rel. 37, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE NADH-ubiquinone oxidoreductase chain 4L (EC 1.6.5.3).
GN MTND4L OR ND4L OR NADH4L.
OS Sus scrofa (Pig).
OG Mitochondrion.
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
OX NCBI_TaxID=9823;
RN [1]
RP SEQUENCE FROM N.A.
RC MEDLINE=96404150; PubMed=9732457;
RA Ursing B.M., Arnason U.;
RA "The complete mitochondrial DNA sequence of the pig (Sus scrofa).";
RL J. Mol. Evol. 47:302-306(1998).
[2]
RP SEQUENCE FROM N.A.
RC STRAIN=Landrace; TISSUE=Heart;
```

```
RX MEDLINE=99365306; PubMed=10433971;
RA Lin C.S., Sun Y.L., Liu C.Y., Yang P.C., Chang L.C., Cheng I.C.,
RA Mao S.J.T., Huang M.C.;
RT "Complete nucleotide sequence of pig (Sus scrofa) mitochondrial genome
RL Gene 236107-114(1999).
CC -1- CATALYTIC ACTIVITY: NADH + ubiquinone = NAD(+) + ubiquinol.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL: AJ002189; CAA05240.1; -.
CC EMBL: AF034253; AAD34193.1; -.
CC PIR: T10980; T10980.
CC InterPro: IPR001133; Oxidored 4L.
CC Pfam: PF00420; oxidored_g2; 1.
CC ProDom: PD000359; Oxidred4L; 1.
CC Oxidoreductase; NAD; Ubiquinone; Mitochondrion.
SQ SEQUENCE 98 AA; 10824 MW; 828C7F511A4E3AFA CRC64;

Query Match 71.0%; Score 22; DB 1; Length 98;
Best Local Similarity 50.0%; Pred. No. 1.7e+02;
Matches 3; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

OY 2 LNVLMA 7
Db 6 MNIIIMA 11

RESULT 14
NULM SHEEP
ID NULM_SHEEP STANDARD; PRT; 98 AA.
AC 078754;
DT 15-DEC-1998 (Rel. 37, Created)
DT 15-DEC-1998 (Rel. 37, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE NADH-ubiquinone oxidoreductase chain 4L (EC 1.6.5.3).
GN MTND4L OR ND4L.
OS Ovis aries (Sheep).
OG Mitochondrion.
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
OC Bovidae; Caprinae; Ovis.
OX NCBI_TaxID=9940;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Merinolandschaf; TISSUE=Liver;
RX MEDLINE=98440761; PubMed=9767689;
RA Hendlender S., Lewalski H., Wassmuth R., Janke A.;
RT "The complete mitochondrial DNA sequence of the domestic sheep (Ovis
RL J. Mol. Evol. 47:441-448(1998).
CC -1- CATALYTIC ACTIVITY: NADH + ubiquinone = NAD(+) + ubiquinol.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL: AF010406; AAD10103.1; -.
CC PIR: T11058.
CC InterPro: IPR001133; Oxidored 4L.
CC Pfam: PF00420; oxidored_g2; 1.
CC
```

```
DR ProDom; PD000359; Oxidred4L; 1.
SQ SEQUENCE 98 AA; 10837 MW; EC0C943C752F8691 CRC64;

Query Match 71.0%; Score 22; DB 1; Length 98;
Best Local Similarity 50.0%; Pred. No. 1.7e+02;
Matches 3; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

OY 2 LNVLMA 7
Db 6 MNIIIMA 11

RESULT 15
YCU5 CAEL
ID YCU5_CAEL STANDARD; PRT; 57 AA.
AC Q22702;
DT 15-JUL-1998 (Rel. 36, Created)
DT 15-JUL-1998 (Rel. 36, Last sequence update)
DT 15-JUL-1998 (Rel. 36, Last annotation update)
DE Hypothetical 6.3 kDa protein T23P2.5 in chromosome X.
GN T23P2.5.
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
OC Rhabditidae; Peloderinae; Caenorhabditis.
OX NCBI_TaxID=6239;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Bristol N2;
RA Du Z.;
RL Submitted (NOV-1995) to the EMBL/GenBank/DBJ databases.
CC -1- SUBCELLULAR LOCATION: Integral membrane protein (Potential).
CC -1- SIMILARITY: BELONGS TO THE UPF0057 (PMP3) FAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL: U39649; AAA80387.1; -.
CC PIR: T16930; T16930.
CC WormPep; T23P2.5; CE05000.
CC InterPro: IPR000612; UPF0057.
CC Pfam: PF01679; UPF0057; 1.
CC PROSITE; PS01309; UPF0057; 1.
CC Hypothetical protein; Transmembrane.
FT TRANSMEM 3 23 POTENTIAL.
FT TRANSMEM 36 56 POTENTIAL.
SQ SEQUENCE 57 AA; 6275 MW; 2A2554626D3F0490 CRC64;

Query Match 67.7%; Score 21; DB 1; Length 57;
Best Local Similarity 50.0%; Pred. No. 1.7e+02;
Matches 3; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

OY 1 ALNVLAM 6
Db 33 AINILL 38

RESULT 16
Y697 ARCFU
ID Y697_ARCFU STANDARD; PRT; 61 AA.
AC Q29561;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Hypothetical protein AF0697.
GN AF0697.
OS Archaeoglobus fulgidus.
OC Archaea; Euryarchaeota; Archaeoglobi; Archaeoglobales;
```

```

OC Archaeoglobaceae; Archaeoglobus.
OX NCBI_TaxID=2234;
RN [1]
RC SEQUENCE FROM N.A.
RX STRAIN=VC-16 / DSM 4304 / ATCC 49558;
RX MEDLINE=98049343; PubMed=9389475;
RA Klenk H.-P., Clayton R.A., Tomb J.-F., White O., Nelson K.E.,
RA Ketchum K.A., Dodson R.J., Winn M., Hickey E.K., Peterson J.D.,
RA Richardson D.L., Karlavage A.R., Graham D.E., Kyriakides N.C.,
RA Fleischmann R.D., Quackenbush J., Lee N.H., Sutton G.G., Gill S.,
RA Kirkness E.F., Dougherty B.A., McKenney K., Adams M.D., Loftus B.,
RA Peterson S., Reich C.I., McNeil L.K., Badger J.H., Gloddeback T.,
RA Overbeek R., Gocayne J.D., Weidman J.F., McDonald L., Uterback T.,
RA Cotton M.D., Spriggs T., Artiach P., Kaine B.P., Sykes S.M.,
RA Sadow P.W., D'Andrea K.P., Bowman C., Fujii C., Garland S.A.,
RA Mason T.M., Olsen G.J., Fraser C.M., Smith H.O., Woese C.R.,
RA Venter J.C.;
RT "The complete genome sequence of the hyperthermophilic, sulphate-
RT reducing archaeon Archaeoglobus fulgidus."
RL Nature 390:364-370(1997).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; AE001056; AAB90542.1; -.
DR PIR; A69337; A69337.
DR TIGR; AF0697; -.
DR TRANSMEM 4 23 POTENTIAL.
KW Hypothetical protein; Transmembrane; Complete proteome.
FT
SQ SEQUENCE 61 AA; 7060 MW; 13B07F57570DF60 CRC64;

Query Match 67.7%; Score 21; DB 1; Length 61;
Best Local Similarity 66.7%; Pred. No. 1.9e+02;
Matches 4; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ALNVLM 6
Db 17 ALNLLL 22

RESULT 17
SPHA HUMAN
ID SPHA_HUMAN STANDARD; PRT; 63 AA.
AC Q15513;
DT 30-MAY-2000 (Rel. 39, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Protein SPHAR (S-phase response protein).
GN SPHAR.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RC SEQUENCE FROM N.A.
RX MEDLINE=95098005; PubMed=7799938;
RA Digweed M., Guenther U., Schneider R., Seyschab H., Friedl R.,
RA Sperling K.;
RT "Irreversible repression of DNA synthesis in Fanconi anemia cells is
RT alleviated by the product of a novel cyclin-related gene."
RL Mol. Cell. Biol. 15:305-314(1995).
RN [2]
RC SEQUENCE FROM N.A.
RP Hall R.;
RA Submitted (NOV-2000) to the EMBL/GenBank/DBJ databases.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -

```

```

CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; X82554; CAAS7901.1; -.
DR EMBL; AL117350; CAC17573.1; -.
DR PIR; I38244; I38244.
DR GO; GO:0006260; P:DNA replication; TAS.
DR GO; GO:0000074; P:regulation of cell cycle; TAS.
SQ SEQUENCE 63 AA; 7515 MW; 81CAB20A4E39C4AP CRC64;

Query Match 67.7%; Score 21; DB 1; Length 63;
Best Local Similarity 80.0%; Pred. No. 1.9e+02;
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ALNVL 5
Db 22 ALNIL 26

RESULT 18
RUXA PYRAB
ID RUXA_PYRAB STANDARD; PRT; 75 AA.
AC Q30V08;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE Putative snRNP Sm-like protein.
GN PYRAB06500 OR PAB8160.
OS Pyrococcus abyssi.
OC Archaea; Euryarchaeota; Thermococci; Thermococcales; Thermococcaceae;
OX Pyrococcus.
OX NCBI_TaxID=29292;
RN [1]
RC SEQUENCE FROM N.A.
RX STRAIN=GB5 / Orsay;
RX PubMed=12622808;
RA Cohen G.N., Barbe V., Flament D., Galperin M., Heilig R., Lecompte O.,
RA Roch O., Prieur D., Querellou J., Ripp R., Thierry J.-C.,
RA Van der Oost J., Weissenbach J., Zivanovic Y., Forterre P.;
RT "An integrated analysis of the genome of the hyperthermophilic
RT archaeon Pyrococcus abyssi."
RL Mol. Microbiol. 47:1495-1512(2003).
CC -I- SIMILARITY: Belongs to the snRNP Sm proteins family.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; AJ248285; CAB49563.1; -.
DR PIR; B75106; B75106.
DR PDB; 1M8V; 11-FEB-03.
DR HAMAP; MF_00257; -.
DR InterPro; IPR006649; snRNP.
DR InterPro; IPR001163; snRNP_Sm.
DR Pfam; PF01423; Sm; 1.
DR ProDom; PD020287; snRNP; 1.
DR SMART; SM00651; Sm; 1.
KW Hypothetical protein; Ribonucleoprotein; Complete proteome;
KW 3D-structure.
SQ SEQUENCE 75 AA; 8489 MW; CB08295C82D03F1E CRC64;

Query Match 67.7%; Score 21; DB 1; Length 75;
Best Local Similarity 66.7%; Pred. No. 2.3e+02;
Matches 4; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Qy 2 LNVLMA 7

```

```

Db      |||:|:|
      38 LNVVLA 43

RESULT 19
RUXX_PVRHO
ID_RUXX_PVRHO STANDARD; PRT; 75 AA.
AC O74016;
DT 15-JUL-1999 (Rel. 38, Created)
DT 15-JUL-1999 (Rel. 38, Last sequence update)
DE 15-SEP-2003 (Rel. 42, Last annotation update)
DE Putative snRNP Sm-like protein.
GN PH1518.2 OR PHS042.
OS Pyrococcus horikoshii.
OC Archaea; Euryarchaeota; Thermococci; Thermococcales; Thermococcaceae;
OC Pyrococcus.
OX NCBI_TaxID=53953;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=OT3;
RA KAWARABAYASI Y., Sawada M., Horikawa H., Haikawa Y., Hino Y.,
RA Yamamoto S., Sekine M., Baba S.-I., Kousugi H., Hosoyama A., Nagai Y.,
RA Sakai M., Ogura K., Otsuka R., Nakazawa H., Takamiya M., Ohfuku Y.,
RA Funahashi T., Tanaka T., Kudoh Y., Yamazaki J., Kushida N., Oguchi A.,
RA Aoki K.-I., Yoshizawa T., Nakamura Y., Robb F.T., Horikoshi K.,
RA Masuchi Y., Shizuya H., Kikuchi H.;
RT "Complete sequence and gene organization of the genome of a hyper-
thermophilic archaeobacterium, Pyrococcus horikoshii OT3.";
RL DNA Res. 5:55-76(1998).
CC -1- SIMILARITY: Belongs to the snRNP Sm proteins family.
-----
This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (see http://www.isb-sib.ch/announcement/
or send an email to license@isb-sib.ch).
-----
EMBL: AP000006; BAA30628.1; -
PIR: D71028; D71028.
DR HAMAP; MF_00257; -; 1.
DR InterPro; IPR006649; snRNP.
DR Pfam; PF01423; Sm; 1.
DR ProDom; PD020287; snRNP; 1.
DR SMART; SM00651; Sm; 1.
KW Hypothetical protein; Ribonucleoprotein; Complete-proteome.
SQ SEQUENCE 75 AA; 8447 MW; CB082950C4CF3F1E CRC64;

Query Match 67.7%; Score 21; DB 1; Length 75;
Best Local Similarity 66.7%; Pred. No. 2.3e+02;
Matches 4; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 2 LNVVLA 7
DB 38 LNVVLA 43

RESULT 20
RUXX_PVRFU
ID_RUXX_PVRFU STANDARD; PRT; 76 AA.
AC Q8U0F4;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DE 28-FEB-2003 (Rel. 41, Last annotation update)
DE Putative snRNP Sm-like protein.
GN PF1542.
OS Pyrococcus furiosus.
OC Archaea; Euryarchaeota; Thermococci; Thermococcales; Thermococcaceae;
OC Pyrococcus.
OX NCBI_TaxID=2261;

```

```

RN RP SEQUENCE FROM N.A.
RC STRAIN=Vc1 / DSM 3638 / ATCC 43587 / JCM 8422;
RA Weiss R.B., Dunn D.M., Robb F.T., Brown J.R.;
RT "The complete sequence of the Pyrococcus furiosus genome.";
RL Submitted (FEB-2002) to the EMBL/GenBank/DBJ databases.
CC -1- SIMILARITY: Belongs to the snRNP Sm proteins family.
-----
This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (see http://www.isb-sib.ch/announcement/
or send an email to license@isb-sib.ch).
-----
EMBL: AE010255; AAL81666.1; -
HAMAP; MF_00257; -; 1.
DR InterPro; IPR006649; snRNP.
DR InterPro; IPR001163; snRNP_Sm.
DR Pfam; PF01423; Sm; 1.
DR ProDom; PD020287; snRNP; 1.
DR SMART; SM00651; Sm; 1.
KW Hypothetical protein; Ribonucleoprotein; Complete proteome.
SQ SEQUENCE 76 AA; 8443 MW; 532A3FF95C434BE1 CRC64;

Query Match 67.7%; Score 21; DB 1; Length 76;
Best Local Similarity 66.7%; Pred. No. 2.4e+02;
Matches 4; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 2 LNVVLA 7
DB 38 LNVVLA 43

RESULT 21
EF1B_METTH
ID_EF1B_METTH STANDARD; PRT; 89 AA.
AC O27734;
DT 15-JUL-1998 (Rel. 36, Created)
DT 15-JUL-1998 (Rel. 36, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Elongation factor 1-beta (EF-1-beta) (aEF-1beta).
GN EF1B OR MTH1699.
OS Methanobacterium thermoautotrophicum.
OC Archaea; Euryarchaeota; Methanobacteria; Methanobacteriales;
OC Methanobacteriacea; Methanothermobacter.
OX NCBI_TaxID=187420;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Delta H;
RX MEDLINE=98037514; PubMed=93711463;
RA Smith D.R., Doucette-Stamm L.A., Deloughery C., Lee H.-M., Dubois J.,
RA Aldredge T., Bashirzadeh R., Blakely D., Cook R., Gilbert K.,
RA Harrison D., Hoang L., Keagle P., Lum W., Pothier B., Qiu D.,
RA Spadafora R., Vicare R., Wang Y., Wierzbowski J., Gibson R.,
RA Jiwani N., Caruso A., Bush D., Safer H., Patwell D., Prabhakar S.,
RA McDougall S., Shimer G., Goyal A., Pietrowski S., Church G.M.,
RA Daniels C.J., Mao J.-I., Rice P., Noelling J., Reeve J.N.;
RT "Complete genome sequence of Methanobacterium thermoautotrophicum
deltaH: functional analysis and comparative genomics.";
RL J. Bacteriol. 179:7135-7155(1997).
RN [2]
RP STRUCTURE BY NMR.
RX MEDLINE=20414184; PubMed=10959626;
RA Kozlov G., Ekliel I., Beglova N., Yee A., Dharamsi A., Engel A.,
RA Siddiqui N., Nong A., Gehring K.;
RT "Rapid fold and structure determination of the archaeal translation
elongation factor 1beta from Methanobacterium thermoautotrophicum.";
RL J. Biomol. NMR 17:187-194(2000).
CC -1- FUNCTION: PROMOTES THE REGENERATION OF GDP FOR GTP IN EF-1-ALPHA/GDP,
CC THUS ALLOWING THE REGENERATION OF EF-1-ALPHA/GTP THAT COULD THEN
CC BE USED TO FORM THE TERNARY COMPLEX EF-1-ALPHA/GTP/AATRNA.

```



CC -1- MISCELLANEOUS: BINDS CALCIUM.  
CC -1- SIMILARITY: BELONGS TO THE EF-1-BETA/EF-1-DELTA FAMILY.  
CC  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC  
CC EMBL; AE000927; A0886171.1; -.  
DR PIR; B69094.  
DR PDB; 1GH8; 13-DEC-00.  
DR HAMAP; MF\_00043; -; 1.  
DR InterPro; IPR004542; aEF-1\_beta.  
DR InterPro; IPR001326; EFL\_BD.  
DR Pfam; PF00736; EFLBD; 1.  
DR TIGRFAMs; TIGR00489; aEF-1\_beta; 1.  
KW Elongation factor; Protein biosynthesis; Calcium-binding;  
KW 3D-structure; Complete proteome.  
FT STRAND 4 12  
FT HELIX 19 29  
FT TURN 32 33  
FT STRAND 35 36  
FT STRAND 41 43  
FT STRAND 49 57  
FT STRAND 62 65  
FT HELIX 66 71  
FT HELIX 72 73  
FT TURN 78 87  
FT STRAND 78 87  
SQ SEQUENCE 89 AA; 9532 MW; 17CC49327D3B773D CRC64;  
  
Query Match 67.7%; Score 21; DB 1; Length 89;  
Best Local Similarity 66.7%; Pred. No. 2.8e+02;  
Matches 4; Conservative 2; Mismatches 0; Indels 0; Gaps 0;  
  
QY 1 ALNVLMA 6  
|||:|  
Db 50 ALNVMV 55  
  
RESULT 22  
IAPP OCTDE STANDARD; PRT; 91 AA.  
AC P22889;  
DT 01-AUG-1991 (Rel. 19, Created)  
DT 01-AUG-1991 (Rel. 19, Last sequence update)  
DT 28-FEB-2003 (Rel. 41, Last annotation update)  
DE Islet amyloid polypeptide precursor (Amylin).  
GN IAPP.  
OS Octodon degus (Degu).  
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Rodentia; Hystriognathi; Octodontidae; Octodon.  
OX NCBI\_TaxID=10160;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE-91155952; PubMed-2293024;  
RA Nishi M., Steiner D.F.;  
RT "Cloning of complementary DNAs encoding islet amyloid polypeptide,  
RT insulin, and glucagon precursors from a New World rodent, the degu,  
RT Octodon degus.";  
RL Mol. Endocrinol. 4:1192-1198(1990).  
CC  
CC -1- FUNCTION: SELECTIVELY INHIBITS INSULIN-STIMULATED GLUCOSE  
CC UTILIZATION AND GLUCOGEN DEPOSITION IN MUSCLE, WHILE NOT AFFECTING  
CC ADIPOCYTE GLUCOSE METABOLISM.  
CC -1- SUBCELLULAR LOCATION: Secreted.  
CC -1- SIMILARITY: BELONGS TO THE CALCITONIN FAMILY.  
CC  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way

CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC  
CC EMBL; M57669; AAA40589.1; -.  
DR PIR; A36118; A36118.  
DR InterPro; IPR000443; Amylin.  
DR InterPro; IPR001693; Calcitonin-like.  
DR InterPro; IPR002163; Calcitonin B.  
DR Pfam; PF00214; Calc\_CGRP\_IAPP; 1.  
DR PRINTS; PR00817; CALCITONINB.  
DR PRINTS; PR00818; ISLETAMYLOID.  
DR SMART; SM00113; CALCITONIN; 1.  
DR PROSITE; PS00258; CALCITONIN; 1.  
KW Hormone; Cleavage on pair of basic residues; Amidation; Amyloid;  
KW Signal.  
FT SIGNAL 1 22 POTENTIAL.  
FT PROPEP 23 34 ISLET AMYLOID POLYPEPTIDE.  
FT PEPTIDE 37 73  
FT PROPEP 77 91  
FT DISULFID 38 43 BY SIMILARITY.  
FT MOD RES 73 73 AMIDATION (G-74 PROVIDE AMIDE GROUP).  
SQ SEQUENCE 91 AA; 9924 MW; 42AB31AEICE9EA99 CRC64;  
  
Query Match 67.7%; Score 21; DB 1; Length 91;  
Best Local Similarity 71.4%; Pred. No. 2.8e+02;  
Matches 5; Conservative 0; Mismatches 2; Indels 0; Gaps 0;  
  
QY 1 ALNVLMA 7  
|||:|  
Db 16 ALNLTXA 22  
  
RESULT 23  
YOHL\_SERMA STANDARD; PRT; 91 AA.  
ID YOHL\_SERMA  
AC P55757;  
DT 01-NOV-1997 (Rel. 35, Created)  
DT 01-NOV-1997 (Rel. 35, Last sequence update)  
DT 16-OCT-2001 (Rel. 40, Last annotation update)  
DE Hypothetical 10.1 kDa protein in bioa 5'region.  
OS Serratia marcescens.  
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;  
OC Enterobacteriaceae; Serratia.  
OX NCBI\_TaxID=615;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=Str41;  
RA Sakurai N., Imai Y., Akatsuka H., Kawai E., Komatsubara S., Tosa T.;  
RL Submitted (AUG-1993) to the EMBL/GenBank/DBJ databases.  
RN [2]  
RP IDENTIFICATION.  
RA Rudd K.E.;  
RL Unpublished observations (MAR-1996).  
CC -1- SIMILARITY: TO E.COLI YAIN.  
CC  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC  
CC EMBL; D17468; -; NOT\_ANNOTATED\_CDS.  
DR InterPro; IPR003735; DUF156.  
DR Pfam; PF02583; DUF156; 1.  
KW Hypothetical protein.  
SQ SEQUENCE 91 AA; 10102 MW; 4E442754797D69F7 CRC64;  
  
Query Match 67.7%; Score 21; DB 1; Length 91;  
Best Local Similarity 71.4%; Pred. No. 2.8e+02;  
Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;



```
QY 1 ALNVLM 7
Db 48 AVANGLMA 54

RESULT 24
NULM_DUGDU STANDARD; PRT; 98 AA.
AC Q8W9M8;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE NADH-ubiquinone oxidoreductase chain 4L (EC 1.6.5.3).
GN MTND4L OR ND4L OR NADH4L.
OS Dugong dugon (Dugong).
OG Mitochondrion.
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Sirenia; Dugongidae; Dugong.
OX NCBI_TaxID=29137;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=22056029; PubMed=12034869;
RA Arnason U., Adegoke J.A., Bodin K., Born E.W., Esa Y.B., Gullberg A.,
RA Nilsson M., Short R.V., Xu X., Janke A.;
RT "Mammalian mitogenomic relationships and the root of the eutherian
tree."
RL Proc. Natl. Acad. Sci. U.S.A. 99:8151-8156(2002).
CC -1- CATALYTIC ACTIVITY: NADH + ubiquinone = NAD(+) + ubiquinol.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; AJ421723; CAD18916.1;
DR InterPro; IPR001133; Oxidored 4L.
DR InterPro; IPR003214; Oxidred4L.
DR Pfam; PF00420; oxidored_q2; 1.
DR ProDom; PD000359; Oxidred4L; 1.
KW Oxidoreductase; NAD; Ubiquinone; Mitochondrion.
SQ SEQUENCE 98 AA; 10921 MW; CCE86F1849FC007C CRC64;

Query Match 67.7%; Score 21; DB 1; Length 98;
Best Local Similarity 50.0%; Pred. No. 3.1e+02;
Matches 3; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 2 LNVLM 7
Db 6 INILLA 11

RESULT 25
NULM_ORNAN STANDARD; PRT; 98 AA.
AC Q36457;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 01-NOV-1997 (Rel. 35, Last annotation update)
DE NADH-ubiquinone oxidoreductase chain 4L (EC 1.6.5.3).
GN MTND4L OR ND4L.
OS Ornithorhynchus anatinus (Duckbill platypus).
OG Mitochondrion.
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Monotremata; Ornithorhynchidae; Ornithorhynchus.
OX NCBI_TaxID=9258;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=97077300; PubMed=8919867;

Query Match 67.7%; Score 21; DB 1; Length 98;
Best Local Similarity 50.0%; Pred. No. 3.1e+02;
Matches 3; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 2 LNVLM 7
Db 6 INILLA 11

RESULT 26
PSAX_SYNVU STANDARD; PRT; 29 AA.
AC P23320;
DT 01-NOV-1991 (Rel. 20, Created)
DT 01-NOV-1991 (Rel. 20, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Photosystem I 4.8 kDa protein (Fragment).
GN PSAX.
OS Synechococcus vulcanus (Thermosynechococcus vulcanus).
OC Bacteria; Cyanobacteria; Chroococcales; Thermosynechococcus.
OX NCBI_TaxID=32053;
RN [1]
RP SEQUENCE.
RX MEDLINE=89338747; PubMed=2503399;
RA Koike H., Ikeuchi M., Hiyama T., Inoue Y.;
RT "Identification of photosystem I components from the cyanobacterium,
RT Synechococcus vulcanus by N-terminal sequencing.";
RL FEBS Lett. 253:257-263(1989).
CC -1- SIMILARITY: BELONGS TO THE PSAX FAMILY.
KW Photosynthesis; Photosystem I.
FT NON TER 29
SQ SEQUENCE 29 AA; 3172 MW; DCF1BBF6418042FF CRC64;

Query Match 64.5%; Score 20; DB 1; Length 29;
Best Local Similarity 57.1%; Pred. No. 1.5e+02;
Matches 4; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 ALNVLM 7
Db 21 AINFLVA 27

RESULT 27
ACPH_UREPA STANDARD; PRT; 77 AA.
AC Q9PPY4;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
```

```
RA Janke A., Gemmell N., Feldmaier-Fuchs G., von Haeseler A.,
RA Paabo S.;
RT "The mitochondrial genome of a monotreme -- the platypus
RT (Ornithorhynchus anatinus).";
RL J. Mol. Evol. 42:153-159(1996).
CC -1- CATALYTIC ACTIVITY: NADH + ubiquinone = NAD(+) + ubiquinol.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; X83427; CAA58463.1;
DR PIR; A58889; A58889.
DR InterPro; IPR001133; Oxidored 4L.
DR InterPro; IPR003214; Oxidred4L.
DR Pfam; PF00420; oxidored_q2; 1.
DR ProDom; PD000359; Oxidred4L; 1.
KW Oxidoreductase; NAD; Ubiquinone; Mitochondrion.
SQ SEQUENCE 98 AA; 10852 MW; 30F80723A3D36090 CRC64;

Query Match 67.7%; Score 21; DB 1; Length 98;
Best Local Similarity 66.7%; Pred. No. 3.1e+02;
Matches 4; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 2 LNVLM 7
Db 6 INILLA 11

RESULT 26
PSAX_SYNVU STANDARD; PRT; 29 AA.
AC P23320;
DT 01-NOV-1991 (Rel. 20, Created)
DT 01-NOV-1991 (Rel. 20, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Photosystem I 4.8 kDa protein (Fragment).
GN PSAX.
OS Synechococcus vulcanus (Thermosynechococcus vulcanus).
OC Bacteria; Cyanobacteria; Chroococcales; Thermosynechococcus.
OX NCBI_TaxID=32053;
RN [1]
RP SEQUENCE.
RX MEDLINE=89338747; PubMed=2503399;
RA Koike H., Ikeuchi M., Hiyama T., Inoue Y.;
RT "Identification of photosystem I components from the cyanobacterium,
RT Synechococcus vulcanus by N-terminal sequencing.";
RL FEBS Lett. 253:257-263(1989).
CC -1- SIMILARITY: BELONGS TO THE PSAX FAMILY.
KW Photosynthesis; Photosystem I.
FT NON TER 29
SQ SEQUENCE 29 AA; 3172 MW; DCF1BBF6418042FF CRC64;

Query Match 64.5%; Score 20; DB 1; Length 29;
Best Local Similarity 57.1%; Pred. No. 1.5e+02;
Matches 4; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 ALNVLM 7
Db 21 AINFLVA 27

RESULT 27
ACPH_UREPA STANDARD; PRT; 77 AA.
AC Q9PPY4;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
```

DE Acyl carrier protein homolog (ACP).  
 GN UUS06.  
 OS Ureaplasma parvum (Ureaplasma urealyticum biotype 1).  
 OC Bacteria; Firmicutes; Mollicutes; Mycoplasmataceae; Ureaplasma.  
 OX NCBI\_TaxID=134821;  
 RN [1]  
 RN SEQUENCE FROM N.A.  
 RC STRAIN=Serovar 3;  
 RX MEDLINE=20500219; PubMed=11048724;  
 RA Glass J.I.; Lefkowitz E.J.; Glass J.S.; Heiner C.R.; Chen E.Y.;  
 RA Cassell G.H.;  
 RT "The complete sequence of the mucosal pathogen Ureaplasma  
 urealyticum.";  
 RL Nature 407:757-762(2000).  
 CC -1- FUNCTION: Carrier of the growing fatty acid chain in fatty acid  
 biosynthesis (By similarity).  
 CC -1- PATHWAY: De novo fatty acid biosynthesis.  
 CC -1- PTM: 4'-Phosphopantetheine is transferred from CoA to a specific  
 serine of the apo-ACP-like protein (potential).  
 CC -1- SIMILARITY: Contains 1 acyl carrier domain.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 the European Bioinformatics Institute. There are no restrictions on its  
 use by non-profit institutions as long as its content is in no way  
 modified and this statement is not removed. Usage by and for commercial  
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL; AEO02149; AAF30918.1; -  
 DR InterPro; IPR006163; Pp bind.  
 DR InterPro; IPR006162; Ppantne attach.  
 DR Pfam; PF00550; Pp-binding; 1.  
 DR PROSITE; PS50075; ACP\_DOMAIN; 1.  
 DR PROSITE; PS00012; PHOSPHOPANTHEINE; 1.  
 DR PROSITE; PS00012; PHOSPHOPANTHEINE; 1.  
 KW Hypothetical protein; Fatty acid biosynthesis; Phosphopantetheine;  
 KW Complete proteome.  
 FT BINDING 36 36 PHOSPHOPANTHEINE (BY SIMILARITY).  
 SQ SEQUENCE 77 AA; 8750 MW; 3CDB655FBFA968C6 CRC64;  
 -----  
 Query Match 64.5%; Score 20; DB 1; Length 77;  
 Best Local Similarity 50.0%; Pred. No. 4.2e+02;  
 Matches 3; Conservative 3; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 ALNVLIM 6  
 |.:|  
 Db 39 ANVLIM 44  
 -----  
 RESULT 28  
 RS20 BORBU  
 ID RS20 BORBU STANDARD; PRT; 85 AA.  
 AC P49334; O51249;  
 DT 01-FEB-1996 (Rel. 33, Created)  
 DT 01-FEB-1996 (Rel. 33, Last sequence update)  
 DT 28-FEB-2003 (Rel. 41, Last annotation update)  
 DE 30S ribosomal protein S20.  
 GN RPST OR B80233.  
 OS Borrelia burgdorferi (Lyme disease spirochete).  
 OC Bacteria; Spirochaetes; Spirochaetales; Spirochaetaceae; Borrelia.  
 OX NCBI\_TaxID=139;  
 RN [1]  
 RN SEQUENCE FROM N.A.  
 RA Tilly K.;  
 RL Submitted (SEP-1995) to the EMBL/GenBank/DBJ databases.  
 RN [2]  
 RN SEQUENCE FROM N.A.  
 RC STRAIN=ATCC 35210 / B31;  
 RX MEDLINE=98065943; PubMed=9403685;  
 RA Fraser C.M.; Casjens S.; Huang W.M.; Sutton G.G.; Clayton R.A.;  
 RA Lathigra R.; White O.; Ketchum K.A.; Dodson R.; Hickey E.K.; Gwinn M.;  
 RA Dougherty B.; Tomb J.-F.; Fleischmann R.D.; Richardson D.;  
 RA Peterson J.; Kerlavage A.R.; Quackenbush J.; Salzberg S.; Hanson M.

RA van Vugt R.; Palmer N.; Adams M.D.; Gocayne J.D.; Weidman J.;  
 RA Uterback T.; Watthey L.; McDonald L.; Artiach P.; Bowman C.;  
 RA Garland S.; Fujii C.; Cotton M.D.; Horst K.; Roberts K.; Hutch B.;  
 RA Smith H.O.; Venter J.C.;  
 RT "Genomic sequence of a Lyme disease spirochaete, Borrelia  
 burgdorferi.";  
 RL Nature 390:580-586(1997).  
 CC -1- FUNCTION: Binds directly to 16S ribosomal RNA (By similarity).  
 CC -1- SIMILARITY: BELONGS TO THE S20P FAMILY OF RIBOSOMAL PROTEINS.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 the European Bioinformatics Institute. There are no restrictions on its  
 use by non-profit institutions as long as its content is in no way  
 modified and this statement is not removed. Usage by and for commercial  
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL; U35673; AAB41459.1; -  
 DR EMBL; AEO01133; AAC66616.1; ALT\_INIT.  
 DR TIGR; BB0233;  
 DR HAMAP; MF\_00500; -; 1.  
 DR InterPro; IPR002583; Ribosomal\_S20p.  
 DR Pfam; PF01649; Ribosomal\_S20p; 1.  
 DR ProDom; PD004231; Ribosomal\_S20p; 1.  
 KW Ribosomal protein; rRNA-binding; Complete proteome.  
 SQ SEQUENCE 85 AA; 9947 MW; EBAC874CD12397FA CRC64;  
 -----  
 Query Match 64.5%; Score 20; DB 1; Length 85;  
 Best Local Similarity 80.0%; Pred. No. 4.7e+02;  
 Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;  
 QY 2 LNVLLM 6  
 |.:|  
 Db 79 LNVLL 83  
 -----  
 RESULT 29  
 CYC6 SYNLI  
 ID CYC6 SYNLI STANDARD; PRT; 87 AA.  
 AC P00114;  
 DT 21-JUL-1986 (Rel. 01, Created)  
 DT 21-JUL-1986 (Rel. 01, Last sequence update)  
 DT 28-FEB-2003 (Rel. 41, Last annotation update)  
 DE Cytochrome c6 (Soluble cytochrome f) (Cytochrome c553) (Cytochrome c-  
 553).  
 GN PETJ.  
 OS Synechococcus lividus.  
 OC Bacteria; Cyanobacteria; Chroococcales; Synechococcus.  
 OX NCBI\_TaxID=33070;  
 RN [1]  
 RN SEQUENCE.  
 RA Borden D.; Margolish E.;  
 RL Submitted (DEC-1979) to the PIR data bank.  
 CC -1- FUNCTION: Functions as an electron carrier between membrane-bound  
 cytochrome b6f and photosystem I in oxygenic photosynthesis (By  
 similarity).  
 CC -1- SUBUNIT: Monomer (By similarity).  
 CC -1- SUBCELLULAR LOCATION: Cytoplasmic (Potential).  
 CC -1- PTM: Binds one heme group per molecule.  
 CC -1- SIMILARITY: Belongs to the cytochrome c family. PetJ subfamily.  
 DR PIR; A00106; CCVC6L.  
 DR HSP; P56534; 1C6S.  
 DR HAMAP; MF\_00594; -; 1.  
 DR InterPro; IPR003088; Cyt\_C1.  
 DR InterPro; IPR002329; Cyt\_C1C.  
 DR Pfam; PF000345; CytC heme bind.  
 DR PRINTS; PR00605; CYTOCHROME\_C; 1.  
 DR PROSITE; PS00130; CYTOCHROME\_C; 1.  
 KW Electron transport; Photosynthesis; Heme.  
 FT BINDING 14 14 HEME (COVALENT).  
 FT BINDING 17 17

FT METAL 18 18 IRON (HEME AXIAL LIGAND).  
 FT METAL 58 58 IRON (HEME AXIAL LIGAND).  
 SQ SEQUENCE 87 AA; 9129 MW; 37713EC6405EBEE CRC64;  
 Query Match 64.5%; Score 20; DB 1; Length 87;  
 Best Local Similarity 80.0%; Pred. No. 4.8e+02;  
 Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;  
 Qy 3 NVLMA 7  
 Db 23 NVVMA 27  
 RESULT 30  
 CYC6 SYNFP3 STANDARD; PRT; 87 AA.  
 AC P00115;  
 DT 21-JUL-1986 (Rel. 01, Created)  
 DT 21-JUL-1986 (Rel. 01, Last sequence update)  
 DT 28-FEB-2003 (Rel. 41, Last annotation update)  
 DE Cytochrome c6 (Soluble cytochrome f) (Cytochrome c553) (Cytochrome c-553) (Cytochrome c-552).  
 GN PETJ.  
 OS Synechococcus sp. (strain PCC 6312 / ATCC 27167).  
 OC Bacteria; Cyanobacteria; Chroococcales; Synechococcus.  
 OX NCBI\_TaxID=195253;  
 RN [1]  
 RP SEQUENCE.  
 RX MEDLINE=80068924; PubMed=228936;  
 RA Aitken A.;  
 RT "Purification and primary structure of cytochrome c-552 from the cyanobacterium, Synechococcus PCC 6312.";  
 RL Eur. J. Biochem. 101:297-308(1979).  
 CC -1- FUNCTION: Functions as an electron carrier between membrane-bound cytochrome b6f and photosystem I in oxygenic photosynthesis.  
 CC -1- SUBUNIT: Monomer (By similarity).  
 CC -1- SUBCELLULAR LOCATION: Cytoplasmic (Potential).  
 CC -1- PFM: Binds one heme group per molecule.  
 CC -1- SIMILARITY: Belongs to the cytochrome c family. PetJ subfamily.  
 DR HSSP; P56534; 1C6S.  
 DR HAMAP; MF 00594; -; 1.  
 DR InterPro; IPR003088; Cyt-CI.  
 DR InterPro; IPR002329; Cyt-CIC.  
 DR Pfam; PF00034; Cytochrome c; 1.  
 DR PRINTS; PR00605; CYTOCHROME C; 1.  
 DR PROSITE; PS00190; CYTOCHROME C; 1.  
 KW Electron transport; Photosynthesis; Heme.  
 FT BINDING 14 14 HEME (COVALENT).  
 FT BINDING 17 17 HEME (COVALENT).  
 FT METAL 18 18 IRON (HEME AXIAL LIGAND).  
 FT METAL 58 58 IRON (HEME AXIAL LIGAND).  
 SQ SEQUENCE 87 AA; 9098 MW; 41FC9734E3DF83F2 CRC64;  
 Query Match 64.5%; Score 20; DB 1; Length 87;  
 Best Local Similarity 80.0%; Pred. No. 4.8e+02;  
 Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;  
 Qy 3 NVLMA 7  
 Db 23 NVVMA 27  
 RESULT 31  
 TRP6 BOVIN STANDARD; PRT; 89 AA.  
 AC Q9MYW0;  
 DT 16-OCT-2001 (Rel. 40, Created)  
 DT 16-OCT-2001 (Rel. 40, Last sequence update)  
 DT 28-FEB-2003 (Rel. 41, Last annotation update)  
 DE Transient receptor potential channel 6 (Fragment).  
 GN TRPC6 OR TRP6.  
 OS Bos taurus (Bovine).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae; Bovidae; Bovinae; Bos.  
 OX NCBI\_TaxID=9913;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Brain;  
 RX MEDLINE=20379038; PubMed=10816590;  
 RA Philipp S., Kretz C., Rautmann J., Himmerkus N., Schroth G., Kretz O., Nastainczyk W., Cavalié A., Hoth M., Flockerzi V.;  
 RT "TRP4 (CCE1) protein is part of native calcium release-activated Ca2+-like channels in adrenal cells.";  
 RL J. Biol. Chem. 275:23965-23972(2000).  
 CC -1- FUNCTION: THOUGHT TO FORM A RECEPTOR-ACTIVATED NON-SELECTIVE CALCIUM PERMEANT CATION CHANNEL. PROBABLY IS OPERATED BY A PHOSPHATIDYLINOSITOL SECOND MESSENGER SYSTEM ACTIVATED BY TYROSINE KINASES OR G-PROTEIN COUPLED RECEPTORS. ACTIVATED BY DIACYLGLYCEROL (DAG) IN A MEMBRANE-DELIMITED FASHION, INDEPENDENTLY OF PROTEIN KINASE C (BY SIMILARITY). SEEMS NOT TO BE ACTIVATED BY INTRACELLULAR CALCIUM STORE DEPLETION (BY SIMILARITY).  
 CC -1- SUBCELLULAR LOCATION: Integral membrane protein (probable).  
 CC -1- SIMILARITY: BELONGS TO THE TRANSIENT RECEPTOR FAMILY. STRPC SUBFAMILY.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to [license@sib-sib.ch](mailto:license@sib-sib.ch)).  
 CC -----  
 DR EMBL; AJ271069; CAC01677.1;  
 DR InterPro; IPR002111; Cat\_Channel\_TrpL.  
 KW Ionic channel; Transmembrane; Ion transport; Calcium channel.  
 FT NON TER 1 1  
 FT TRANSMEM 1 21 POTENTIAL.  
 FT TRANSMEM 68 88 POTENTIAL.  
 FT NON TER 89 89  
 SQ SEQUENCE 89 AA; 10315 MW; 8E1D54F051F1E7ED CRC64;  
 Query Match 64.5%; Score 20; DB 1; Length 89;  
 Best Local Similarity 66.7%; Pred. No. 4.9e+02;  
 Matches 4; Conservative 2; Mismatches 0; Indels 0; Gaps 0;  
 Qy 2 NVLMA 7  
 Db 81 NVMLIA 86  
 RESULT 32  
 ILGI\_CABEL STANDARD; PRT; 91 AA.  
 AC Q18060;  
 DT 15-JUL-1998 (Rel. 36, Created)  
 DT 15-JUL-1998 (Rel. 36, Last sequence update)  
 DT 28-FEB-2003 (Rel. 41, Last annotation update)  
 DE Probable insulin-like peptide gamma-type 1 precursor (Ceinsulin-3).  
 GN INS-11 OR C17C3.4.  
 OS Caenorhabditis elegans.  
 OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea; Rhabditidae; Peloderinae; Caenorhabditis.  
 OX NCBI\_TaxID=6239;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Kawano T.;  
 RT "mRNA for a putative insulin-like peptide of Caenorhabditis elegans.";  
 RL Submitted (SEP-1999) to the EMBL/GenBank/DBJ databases.  
 RN [2]  
 RP SEQUENCE FROM N.A.

RC STRAIN=Bristol N2;  
RA Du Z.;  
RL Submitted (DEC-1995) to the EMBL/GenBank/DBJ databases.  
RN [3]  
RP SMILARITY TO INSULIN.  
RX MEDLINE=98217375; PubMed=9548970;  
RA Duret L., Gueux N., Peitsch M.C., Bairoch A.;  
RT "New insulin-like proteins with atypical disulfide bond pattern  
RT characterized in Caenorhabditis elegans by comparative sequence  
RT analysis and homology modeling";  
RL Genome Res. 8:348-353(1998).  
CC -1- SUBCELLULAR LOCATION: Secreted (Potential).  
CC -1- SIMILARITY: BELONGS TO THE INSULIN/IGF/RELAXIN FAMILY.  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
DR EMBL; AB032258; BAA84470.1; -;  
DR EMBL; U41279; AAK31418.1; -;  
DR PIR; T37327; T37327.  
DR WormPep; C17C3.4; CE04024.  
DR InterPro; IPR004825; Ins/IGF/relax.  
DR PROSITE; PS00262; INSULIN; 1.  
KW Insulin family; Signal.  
FT SIGNAL 1 26  
FT CHAIN 27 91  
FT FT  
FT PEPTIDE 34 58  
FT B-CHAIN-LIKE PEPTIDE (POTENTIAL).  
FT A-CHAIN-LIKE PEPTIDE (POTENTIAL).  
FT DISULFID 37 66  
FT POTENTIAL.  
FT DISULFID 49 79  
FT POTENTIAL.  
FT DISULFID 65 70  
FT POTENTIAL.  
SQ SEQUENCE 91 AA; 10173 MW; 22BF958BF759F254 CRC64;  
  
Query Match 64.5%; Score 20; DB 1; Length 91;  
Best Local Similarity 80.0%; Pred. No. 5e+02;  
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;  
  
Qy 1 ALNVL 5  
Db 44 ALNVM 48  
|||||  
  
RESULT 33  
FLIE AQUAE STANDARD; PRT; 93 AA.  
AC O67242;  
DT 30-MAY-2000 (Rel. 39, Created)  
DT 30-MAY-2000 (Rel. 39, Last sequence update)  
DT 16-OCT-2001 (Rel. 40, Last annotation update)  
DE Flagellar hook-basal body complex protein flie.  
GN FLIE OR AQ 1182.1 OR AQ 1182A.  
OS Aquifex aeolicus.  
OC Bacteria; Aquificae; Aquificales; Aquificaceae; Aquifex.  
OX NCBI\_TaxID=63363;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX STRAIN=VF5;  
RC MEDLINE=98196666; PubMed=9537320;  
RA Deckert G., Warren P.V., Gaasterland T., Young W.G., Lenox A.L.,  
RA Graham D.E., Overbeek R., Sneed M.A., Keller M., Aufay M., Huber R.,  
RA Feldman R.A., Short J.M., Olson G.J., Swanson R.V.;  
RT "The complete genome of the hyperthermophilic bacterium Aquifex  
RT aeolicus";  
RL Nature 392:353-358(1998).  
CC -1- SIMILARITY: BELONGS TO THE FLIE FAMILY.  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC -----

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
DR EMBL; AE000727; AAC07208.1; -;  
DR PIR; G70401; G70401.  
DR InterPro; IPR001624; Flie.  
DR Pfam; PF02049; Flie; 1.  
DR PRINTS; PR01006; FLGHOOKFLIE.  
DR TIGRFAMS; TIGR00205; flie; 1.  
KW Flagella; Complete proteome.  
SQ SEQUENCE 93 AA; 10764 MW; 6180AEALC8FB9C39 CRC64;  
  
Query Match 64.5%; Score 20; DB 1; Length 93;  
Best Local Similarity 66.7%; Pred. No. 5.1e+02;  
Matches 4; Conservative 2; Mismatches 0; Indels 0; Gaps 0;  
  
Qy 1 ALNVL 6  
Db 70 ALNLLI 75  
|||||  
  
RESULT 34  
YGBF ECOLI STANDARD; PRT; 94 AA.  
ID YGBF ECOLI  
AC P45956;  
DT 01-NOV-1995 (Rel. 32, Created)  
DT 30-MAY-2000 (Rel. 39, Last sequence update)  
DT 16-OCT-2001 (Rel. 40, Last annotation update)  
DE Hypothetical protein ygbf.  
GN YGBF OR B2754.  
OS Escherichia coli.  
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;  
OC Enterobacteriaceae; Escherichia.  
OX NCBI\_TaxID=562;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX STRAIN=K12;  
RC MEDLINE=89255132; PubMed=2656660;  
RA Nakata A., Amemura M., Makino K.;  
RT "Unusual nucleotide arrangement with repeated sequences in the  
RT Escherichia coli K-12 chromosome.";  
RL J. Bacteriol. 171:3553-3556(1989).  
RN [2]  
RP SEQUENCE FROM N.A.  
RX STRAIN=K12 / MGI655;  
RC MEDLINE=97426617; PubMed=9278503;  
RA Blattner F.R., Plunkett G. III, Bloch C.A., Perna N.T., Burland V.,  
RA Riley M., Collado-Vides J., Glasner J.D., Rode C.K., Mayhew G.P.,  
RA Gregor J., Davis N.W., Kirkpatrick H.A., Goeden M.A., Rose D.J.,  
RA Mau B., Shao Y.;  
RT "The complete genome sequence of Escherichia coli K-12.";  
RL Science 277:1453-1474(1997).  
RN [3]  
RP IDENTIFICATION.  
RX MEDLINE=96032851; PubMed=7567469;  
RA Borodovsky M., McIninch J., Koonin E.V., Rudd K.E., Medigue C.,  
RA Danchin A.;  
RT "Detection of new genes in a bacterial genome using Markov models for  
RT three gene classes";  
RL Nucleic Acids Res. 23:3554-3562(1995).  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----

```

DR EMBL; M27059; -, NOT ANNOTATED_CDS.
DR EMBL; U29579; AAA69264.1; -.
DR EMBL; AE000359; AAC75796.1; ALT_INIT.
DR EcoGene; EG12845; ygbF.
KW Hypothetical protein; Complete proteome.
SQ SEQUENCE 94 AA; 10518 MW; D1C159D924B477B4 CRC64;

Query Match 64.5%; Score 20; DB 1; Length 94;
Best Local Similarity 80.0%; Pred. No. 5.2e+02;
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 3 NVLMA 7
Db 55 NVWMA 59

RESULT 35
YD82_ARCFU
ID_YD82_ARCFU STANDARD; PRT; 95 AA.
AC O28889;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DE 16-OCT-2001 (Rel. 40, Last annotation update)
DE Hypothetical protein AF1382.
GN AF1382.
OS Archaeoglobus fulgidus.
OC Archaea; Euryarchaeota; Archaeoglobi; Archaeoglobales;
OC Archaeoglobaceae; Archaeoglobus.
OX NCBI_TaxID=2234;
RN [1]
SEQUENCE FROM N.A.
RC STRAIN=VC-16 / DSM 4304 / ATCC 49558;
RX MEDLINE=98049343; PubMed=9389475;
RA Klenk H.-P., Clayton R.A., Tomb J.-F., White O., Nelson K.E.,
RA Ketchum K.A., Dodson R.J., Gwinn M., Hickey E.K., Peterson J.D.,
RA Richardson D.L., Kerlavage A.R., Graham D.E., Kyrpides N.C.,
RA Fleischmann R.D., Quackenbush J., Lee N.H., Sutton G.G., Gill S.,
RA Kirkness E.F., Dougherty B.A., McKenney K., Adams M.D., Loftus B.,
RA Peterson S., Reich C.I., McNeil L.K., Badger J.H., Glodek A., Zhou L.,
RA Overbeek R., Gocayne J.D., Weidman J.F., McDonald L., Uterback T.,
RA Corton M.D., Spriggs T., Artiach P., Kaine B.P., Sykes S.M.,
RA Sadow P.W., D'Andrea K.P., Bowman C., Fujii C., Garland S.A.,
RA Mason T.M., Olsen G.J., Fraser C.M., Smith H.O., Woese C.R.,
RA Venter J.C.;
RT "The complete genome sequence of the hyperthermophilic, sulphate-
reducing archaeon Archaeoglobus fulgidus.";
RL Nature 390:364-370(1997).

-----
This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).
-----
DR EMBL; AE001009; AAB89883.1; -.
DR TIGR; AF1382; -.
KW Hypothetical protein; Complete proteome.
SQ SEQUENCE 95 AA; 11144 MW; E6F5FC3126C1E97A CRC64;

Query Match 64.5%; Score 20; DB 1; Length 95;
Best Local Similarity 66.7%; Pred. No. 5.3e+02;
Matches 4; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 1 ALNVLM 6
Db 14 ALEILM 19

RESULT 36
NULM_DIDMA
ID_NULM_DIDMA STANDARD; PRT; 98 AA.
AC P41307;
DT 01-FEB-1995 (Rel. 31, Created)
DT 01-FEB-1995 (Rel. 31, Last sequence update)
DT 01-NOV-1997 (Rel. 35, Last annotation update)
DE NADH-ubiquinone oxidoreductase chain 4L (EC 1.6.5.3).
GN MTND4L OR ND4L.
OS Didelphis marsupialis virginiana (North American opossum).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Metatheria; Didelphimorphia; Didelphidae; Didelphis.
OX NCBI_TaxID=9267;
RN [1]
SEQUENCE FROM N.A.
RC TISSUE=Liver;
RX MEDLINE=94333786; PubMed=8056314;
RA Janke A., Feldmaier-Fuchs G., Thomas K., von Haeseler A., Paabo S.;
RT "The marsupial mitochondrial genome and the evolution of placental
mammals.";
RL Genetics 137:243-256(1994).
CC -|- CATALYTIC ACTIVITY: NADH + ubiquinone = NAD(+) + ubiquinol.
-----
This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).
-----
DR EMBL; Z29573; CAA82685.1; -.
DR PIR; S47878; S47878.
DR InterPro; IPR001133; Oxidored 4L.
DR InterPro; IPR003214; Oxidred4L.
DR Pfam; PF00420; oxidored_02; 1.
DR ProDom; PD000359; Oxidred4L; 1.
DR Oxidoreductase; NAD; Ubiquinone; Mitochondrion.
SQ SEQUENCE 98 AA; 10815 MW; F988F6026B0A4C59 CRC64;

Query Match 64.5%; Score 20; DB 1; Length 98;
Best Local Similarity 50.0%; Pred. No. 5.4e+02;
Matches 3; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

Qy 2 LNLVMA 7
Db 6 LNIIVA 11

RESULT 37
NULM_MACRO
ID_NULM_MACRO STANDARD; PRT; 98 AA.
AC P92657;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE NADH-ubiquinone oxidoreductase chain 4L (EC 1.6.5.3).
GN MTND4L OR ND4L OR NADH4L.
OS Macropus robustus (Wallaroo) (Euro).
OC Mitochondrion.
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Metatheria; Diprotodontia; Macropodidae; Macropus.
OX NCBI_TaxID=9319;
RN [1]
SEQUENCE FROM N.A.
RX MEDLINE=97188458; PubMed=9037043;
RA Janke A., Xu X., Arnason U.;
RT "The complete mitochondrial genome of the wallaroo (Macropus
robustus) and the phylogenetic relationship among Monotremata,
Marsupialia, and Eutheria.";
RL Proc. Natl. Acad. Sci. U.S.A. 94:1276-1281(1997).
CC -|- CATALYTIC ACTIVITY: NADH + ubiquinone = NAD(+) + ubiquinol.
-----
This SWISS-PROT entry is copyright. It is produced through a collaboration

```

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----

DR EMBL; Y10524; CAA71544.1; --  
 DR PIR; T11436; T11436.  
 DR InterPro; IPR001133; Oxidored 4L.  
 DR Pfam; PF00420; Oxidored G2; 1.  
 DR ProDom; PD000359; Oxidred4L; 1.  
 KW Oxidoreductase; NAD; Ubiquinone; Mitochondrion.  
 SQ SEQUENCE 98 AA; 10717 MW; DC9ALD06DBAC81B7 CRC64;

Query Match 64.5%; Score 20; DB 1; Length 98;  
 Best Local Similarity 57.1%; Pred. No. 5.4e+02;  
 Matches 4; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Qy 1'ALNVLMA 7  
 Db 5 SLNLTMA 11  
 :||: ||  
 :||: ||

RESULT 38  
 IPYR\_PSEAN STANDARD; PRT; 25 AA.  
 AC P80898;  
 DT 28-FEB-2003 (Rel. 41, Created)  
 DT 28-FEB-2003 (Rel. 41, Last sequence update)  
 DT 28-FEB-2003 (Rel. 41, Last annotation update)  
 DE Inorganic pyrophosphatase (EC 3.6.1.1) (Pyrophosphate phospho-  
 DE hydrolase) (ppase) (fragment).  
 GN PPA.  
 OS Pseudanabaena sp. (strain PCC 6901).  
 OC Bacteria; Cyanobacteria; Oscillatoriales; Pseudanabaena.

QX NCBI\_TaxID=47918;  
 RN [1]  
 RP SEQUENCE.  
 RA Gomez R., Serrano A.;  
 RL Submitted (FEB-1997) to the SWISS-PROT data bank.  
 CC -!- FUNCTION: Hydrolyzes PPI generated in anabolic reactions.  
 CC -!- CATALYTIC ACTIVITY: Diphosphate + H(2)O = 2 phosphate.  
 CC -!- COFACTOR: Requires the presence of divalent metal cation.  
 CC -!- MAGNESIUM: Requires the highest activity. Binds 4 divalent cations  
 CC per subunit (By similarity).  
 CC -!- SUBUNIT: Homohexamer (By similarity).  
 CC -!- SUBCELLULAR LOCATION: Cytoplasmic.  
 CC -!- SIMILARITY: Belongs to the Ppase family.  
 DR HAMAP; MF\_00209; -; 1.  
 KW Hydrolase; Magnesium.  
 FT NON\_TER 25  
 SQ SEQUENCE 25 AA; 2630 MW; FB6138F3AED4D43E CRC64;

Query Match 61.3%; Score 19; DB 1; Length 25;  
 Best Local Similarity 80.0%; Pred. No. 2.3e+02;  
 Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;  
 Qy 2 LNVLM 6  
 Db 15 LNVLI 19  
 :||: ||  
 :||: ||

RESULT 39  
 PKL\_DICDI STANDARD; PRT; 33 AA.  
 AC P34101;  
 DT 01-FEB-1994 (Rel. 28, Created)  
 DT 01-FEB-1994 (Rel. 28, Last sequence update)  
 DT 16-OCT-2001 (Rel. 40, Last annotation update)  
 DE Protein kinase 1 (EC 2.7.1.-) (Fragment).  
 GN PKGA.

OS Dictyostelium discoideum (Slime mold).  
 OC Eukaryota; Mycetozoa; Dictyosteliida; Dictyostelium.  
 OX NCBI\_TaxID=44689;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=91142122; PubMed=1996312;  
 RA Haribabu B., Dottin R.P.;  
 RT "Identification of a protein kinase multigene family of Dictyostelium  
 RT discoideum: molecular cloning and expression of a cDNA encoding a  
 RT developmentally regulated protein kinase";  
 RL Proc. Natl. Acad. Sci. U.S.A. 88:1115-1119(1991).  
 CC -!- SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.  
 CC -----

CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----

DR EMBL; M59745; AAA33187.1; --  
 DR DictyDb; DD05046; pkga.  
 DR InterPro; IPR000719; Prot\_kinase.  
 DR InterPro; IPR002290; Ser\_thr\_kinase.  
 DR Pfam; PF00069; pkinase; 1.  
 DR PROSITE; PS00107; PROTEIN\_KINASE\_ATP; PARTIAL.  
 DR PROSITE; PS00108; PROTEIN\_KINASE\_ST; PARTIAL.  
 DR PROSITE; PS00111; PROTEIN\_KINASE\_DOM; 1  
 KW Transferase; Serine/threonine-protein kinase; ATP-binding;  
 KW Phosphorylation.  
 FT NON\_TER 1 1  
 FT MOD\_RES <1 >33 PROTEIN\_KINASE.  
 FT MOD\_RES 31 31 PHOSPHORYLATION (AUTO-) (BY SIMILARITY).  
 FT NON\_TER 33 33  
 SQ SEQUENCE 33 AA; 3682 MW; D4CBB6AF12FA18DE CRC64;

Query Match 61.3%; Score 19; DB 1; Length 33;  
 Best Local Similarity 60.0%; Pred. No. 3.1e+02;  
 Matches 3; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Qy 3 NVLMA 7  
 Db 1 NILLA 5  
 :||: ||  
 :||: ||

RESULT 40  
 PSAX\_ANASP STANDARD; PRT; 43 AA.  
 AC P58566;  
 DT 28-FEB-2003 (Rel. 41, Created)  
 DT 28-FEB-2003 (Rel. 41, Last sequence update)  
 DT 28-FEB-2003 (Rel. 41, Last annotation update)  
 DE Photosystem I 4.8 kDa protein.  
 GN PSAX OR ASRI283.  
 OS Anabaena sp. (strain PCC 7120).  
 OC Bacteria; Cyanobacteria; Nostocales; Nostocaceae; Nostoc.  
 OX NCBI\_TaxID=103690;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=21595285; PubMed=11759840;  
 RA Kaneko T., Nakamura Y., Wolk C.P., Kuritz T., Sasamoto S.,  
 RA Watanabe A., Iriguchi M., Ishikawa A., Kawashima K., Kimura T.,  
 RA Kishida Y., Kohara M., Matsumoto M., Matsuno A., Muraki A.,  
 RA Nakazaki N., Shimpo S., Sugimoto M., Takazawa M., Yamada M.,  
 RA Yasuda M., Tabata S.;  
 RT "Complete genomic sequence of the filamentous nitrogen-fixing  
 RT cyanobacterium Anabaena sp. strain PCC 7120.";  
 RL DNA Res. 8:205-213(2001).  
 CC -!- SIMILARITY: BELONGS TO THE PSAX FAMILY.  
 CC -----

CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -

CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

DR ENBL; AP003585; BAB73240.1; -  
DR PIR; AH1966; AH1966.  
KW Photosystem I; Photosynthesis; Complete proteome.  
FT INIT MET 0  
SQ SEQUENCE 43 AA; 4737 MW; 086732FB67A59FEC CRC64;  
BY SIMILARITY.

Query Match 61.3%; Score 19; DB 1; Length 43;  
Best Local Similarity 57.1%; Pred. No. 4.1e+02;  
Matches 4; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 ALNVLMA 7  
DB 29 AVNFLVA 35

## RESULT 41

ID Y567\_PSEAE STANDARD; PRT; 52 AA.  
AC Q915W9;  
DT 16-OCT-2001 (Rel. 40, Created)  
DT 16-OCT-2001 (Rel. 40, Last sequence update)  
DT 16-OCT-2001 (Rel. 40, Last annotation update)  
DE Hypothetical protein PA0567.  
GN PA0567.

OS Pseudomonas aeruginosa.  
OC Bacteria; Proteobacteria; Gammaproteobacteria; Pseudomonadales;  
OC Pseudomonadaceae; Pseudomonas.

OX NCBI\_TaxID=287;  
RN [1]  
RP SEQUENCE FROM N.A.

RC STRAIN=ATCC 15692 / PA01;  
RX MEDLINE=20437337; PubMed=10984043;  
RA Stover C.K., Pham X.-Q.T., Etwin A.L., Mizoguchi S.D., Warren P.,  
RA Hackey M.J., Brinkman F.S.L., Hufnagle W.O., Kowalik D.J., Lagrou M.,  
RA Garber R.L., Goltry L., Tolentino E., Westbrook-Wadman S., Yuan Y.,  
RA Brody L.L., Coulter S.N., Folger S.N., Kas A., Larbig K., Lim R.M.,  
RA Smith K.A., Spencer D.H., Wong G.K.-S., Wu Z., Paulsen I.T.,  
RA Reizer J., Sailer M.H., Hancock R.E.W., Lory S., Olson M.V.;  
RT "Complete genome sequence of Pseudomonas aeruginosa PA01, an  
RT opportunistic pathogen."  
RL Nature 406:959-964(2000).

CC -1- SUBCELLULAR LOCATION: Integral membrane protein (Potential).  
CC -1- SIMILARITY: BELONGS TO THE UPF0057 (PMP3) FAMILY.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

DR ENBL; AE004493; AAG03956.1; -  
DR PIR; H83573; H83573.  
DR InterPro; IPR000612; UPF0057.  
DR Pfam; PF01679; UPF0057; 1.

DR PROSITE; PS01309; UPF0057; 1.  
KW Hypothetical protein; Transmembrane; Complete proteome.  
FT TRANSMEM 6 26 POTENTIAL.  
FT TRANSMEM 29 49 POTENTIAL.

SQ SEQUENCE 52 AA; 5717 MW; CF552458732A04CC CRC64;

Query Match 61.3%; Score 19; DB 1; Length 52;  
Best Local Similarity 60.0%; Pred. No. 5e+02;  
Matches 3; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 2 LNVLM 6

Db 29 LNTLL 33  
||:|:

## RESULT 42

ID COXE\_BOVIN STANDARD; PRT; 85 AA.  
AC P3182;  
DT 01-JAN-1990 (Rel. 13, Created)  
DT 01-FEB-1996 (Rel. 33, Last sequence update)  
DT 16-OCT-2001 (Rel. 40, Last annotation update)  
DE Cytochrome c oxidase polypeptide Via-liver (EC 1.9.3.1) (SSG).  
GN COX6A1.

OS Bos taurus (Bovine).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;  
OC Bovidae; Bovinae; Bos.

OX NCBI\_TaxID=9913;  
RN [1]  
RP SEQUENCE FROM N.A.

RC TISSUE=Liver;  
RX MEDLINE=92070527; PubMed=1720401;

RA Ewart G.D., Zhang Y.-Z., Capaldi R.A.;

RT "Switching of bovine cytochrome c oxidase subunit Via isoforms in  
RT skeletal muscle during development.";

RL FBS Lett. 292:79-84(1991).  
RN [2]  
RP SEQUENCE OF 1-32.

RC TISSUE=Liver;  
RX MEDLINE=89000697; PubMed=2844245;

RA Yamamura W., Zhang Y.-Z., Takamiya S., Capaldi R.A.;

RT "Tissue-specific differences between heart and liver cytochrome c  
RT oxidase.";

RL Biochemistry 27:4909-4914(1988).  
CC -1- FUNCTION: THIS PROTEIN IS ONE OF THE NUCLEAR-CODED POLYPEPTIDE  
CC CHAINS OF CYTOCHROME C OXIDASE, THE TERMINAL OXIDASE IN  
CC MITOCHONDRIAL ELECTRON TRANSPORT.

CC -1- CATALYTIC ACTIVITY: 4 ferrocyclochrome c + O(2) = 4 ferricytochrome  
CC c + 2 H(2)O.

CC -1- SUBCELLULAR LOCATION: Mitochondrial inner membrane.

CC -1- SIMILARITY: BELONGS TO THE CYTOCHROME C OXIDASE VIA FAMILY.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

DR ENBL; M38520; AAA30437.1; -  
DR PIR; S18314; S18314.

DR InterPro; IPR001349; COX6A.

DR Pfam; PF02046; COX6A; 1.

DR ProDom; PD006036; COX6A; 1.

DR PROSITE; PS01329; COX6A; 1.

KW Oxidoreductase; Inner membrane; Mitochondrion.

FT CONFLICT 19 19 L -> Y (IN REF. 2).

FT CONFLICT 28 28 V -> L (IN REF. 2).

FT CONFLICT 30 30 M -> T (IN REF. 2).

SQ SEQUENCE 85 AA; 9507 MW; 9F493F6979E2E74B CRC64;

Query Match 61.3%; Score 19; DB 1; Length 85;  
Best Local Similarity 80.0%; Pred. No. 8.3e+02;

Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 LNVLM 6

Db 31 LNVFM 35  
|||:

## RESULT 43

C555\_CHLIT

ID C555\_CHLIT STANDARD; PRT; 86 AA.  
 AC P00123;  
 DT 21-JUL-1986 (Rel. 01, Created)  
 DT 21-JUL-1986 (Rel. 01, Last sequence update)  
 DE Cytochrome c-555 (C555)  
 OS Chlorobium limicola f.sp. thiosulfatophilum.  
 OC Bacteria; Chlorobi; Chlorobiales; Chlorobiaceae;  
 OC Chlorobium.  
 OX NCBI\_TaxID=115852;  
 RN [1]  
 RP SEQUENCE.  
 RX MEDLINE=77087089; PubMed=188412;  
 RA van Beunten J., Ambler R.P., Meyer T.E., Kamen M.D., Olson J.M.,  
 RA Shaw E.K.,  
 RT "The amino acid sequences of the cytochromes c-555 from two green  
 RT sulphur bacteria of the genus Chlorobium."  
 RL Biochem. J. 159:757-774 (1976).  
 RN [2]  
 RP X-RAY CRYSTALLOGRAPHY (2-7 ANGSTROMS).  
 RX MEDLINE=78094383; PubMed=202947;  
 RA Korsun Z.R., Saleme F.R.;  
 RT "Structure of cytochrome c555 of Chlorobium thiosulfatophilum:  
 RT primitive low-potential cytochrome c."  
 RL Proc. Natl. Acad. Sci. U.S.A. 74:5244-5247 (1977).  
 CC -1- FUNCTION: This basic c-type monoheme cytochrome has been found  
 CC exclusively in the green photosynthetic bacteria, although its  
 CC role in bacterial photosynthesis is not established. It has an  
 CC unusually low redox potential compared with mitochondrial  
 CC cytochrome c. It is reactive with cytochrome c oxidases but not  
 CC with reductases.  
 CC HSP; P11732; 1CC5.  
 DR InterPro; IPR003088; Cyt CI.  
 DR InterPro; IPR000345; CytC\_heme\_bind.  
 DR Pfam; PF00034; Cytochrome c; 1.  
 DR PROSITE; PS00190; CYTOCHROME C; 1.  
 KW Electron transport; Photosynthesis; Heme.  
 FT BINDING 14 14 HEME (COVALENT).  
 FT BINDING 17 17 HEME (COVALENT).  
 FT METAL 18 18 IRON (HEME AXIAL LIGAND).  
 FT METAL 60 60 IRON (HEME AXIAL LIGAND).  
 SQ SEQUENCE 86 AA; 8780 MW; 0882D21350DB9D4E CRC64;  
  
 Query Match 61.3%; Score 19; DB 1; Length 86;  
 Best Local Similarity 50.0%; Pred. NO. 8.4e+02;  
 Matches 3; Conservative 3; Mismatches 0; Indels 0; Gaps 0;  
  
 QY 2 LNVLMA 7  
 Db 42 MNVWVA 47  
  
 RESULT 44  
 YM17 MARPO STANDARD; PRT; 86 AA.  
 ID C39450;  
 DT 01-OCT-1994 (Rel. 30, Created)  
 DT 01-OCT-1994 (Rel. 30, Last sequence update)  
 DE 16-OCT-2001 (Rel. 40, Last annotation update)  
 DE Hypothetical 10.2 kDa protein in RPS1-NAD4L intergenic region  
 DE (ORF 86A).  
 GN YMF17.  
 OS Marchantia polymorpha (Liverwort).  
 OG Mitochondrion.  
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Marchantiophyta;  
 OC Marchantiopsida; Marchantiidae; Marchantiales; Marchantiaceae;  
 OC Marchantiaceae; Marchantia.  
 OX NCBI\_TaxID=3197;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=92114051; PubMed=1731062;  
 RA Oda K., Yamato K., Ohta E., Nakamura Y., Takemura M., Nozato N.,  
 RA Akashi K., Kanegae T., Ogura Y., Kohchi T., Ohnaya K.;

RT "Gene organization deduced from the complete sequence of liverwort  
 RT Marchantia polymorpha mitochondrial DNA. A primitive form of plant  
 RT mitochondrial genome."  
 RL J. Mol. Biol. 223:1-7 (1992).  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See http://www.isb-sib.ch/announce/  
 CC or send an email to license@isb-sib.ch).  
 CC -----  
 CC EMBL; M68929; AAC09428.1; -.  
 DR PIR; S25987; S25987.  
 KW Mitochondrion; Hypothetical protein.  
 SQ SEQUENCE 86 AA; 10237 MW; 809246B7664A5F17 CRC64;  
  
 Query Match 61.3%; Score 19; DB 1; Length 86;  
 Best Local Similarity 60.0%; Pred. NO. 8.4e+02;  
 Matches 3; Conservative 2; Mismatches 0; Indels 0; Gaps 0;  
  
 QY 2 LNVLMA 6  
 Db 36 LNILL 40  
  
 RESULT 45  
 SYNC\_SACDO STANDARD; PRT; 87 AA.  
 ID P41908;  
 DT 01-NOV-1995 (Rel. 32, Created)  
 DT 01-NOV-1995 (Rel. 32, Last sequence update)  
 DT 28-FEB-2003 (Rel. 41, Last annotation update)  
 DE Putative asparaginyl-tRNA synthetase, cytoplasmic (EC 6.1.1.22)  
 DE (Asparagine--tRNA ligase) (AsnRS) (Fragment).  
 GN DED81.  
 OS Saccharomyces douglasii (Yeast).  
 OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;  
 OC Saccharomycetales; Saccharomycetaceae; Saccharomycetes.  
 OX NCBI\_TaxID=46617;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=94287708; PubMed=8017101;  
 RA Adjiri A., Chanet R., Mezard C., Fabre F.;  
 RT "Sequence comparison of the ARG4 chromosomal regions from the two  
 RT related yeasts, Saccharomyces cerevisiae and Saccharomyces  
 RT douglasii."  
 RL Yeast 10:309-317 (1994).  
 CC -1- CATALYTIC ACTIVITY: ATP + L-asparagine + tRNA(Asn) = AMP +  
 CC diphosphate + L-asparaginyl-tRNA(Asn).  
 CC -1- SUBCELLULAR LOCATION: Cytoplasmic.  
 CC -1- SIMILARITY: Belongs to class-II aminoacyl-tRNA synthetase family.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See http://www.isb-sib.ch/announce/  
 CC or send an email to license@isb-sib.ch).  
 CC -----  
 CC EMBL; X73886; CAA52090.1; -.  
 DR HSP; P04802; IASZ.  
 DR InterPro; IPR004364; tRNA-synt\_2.  
 DR InterPro; IPR002312; tRNA-synt\_asp.  
 DR InterPro; IPR006195; tRNA ligase II.  
 DR Pfam; PF00152; tRNA-synt 2; 1.  
 DR PRINTS; PR01042; TRNASYNTHASP.  
 DR PROSITE; PS50862; AA tRNA LIGASE II; PARTIAL.  
 KW Aminoacyl-tRNA synthetase; Protein biosynthesis; Ligase; ATP-binding.  
 FT NON\_TER 1  
 SQ SEQUENCE 87 AA; 9934 MW; 98174B082C64E78 CRC64;



Query Match 61.3%; Score 19; DB 1; Length 87;  
Best Local Similarity 66.7%; Pred. No. 8.5e+02;  
Matches 4; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 2 LNVLMA 7  
: | | | |  
Db 21 MNELMA 26

Search completed: February 17, 2004, 10:57:05  
Job time : 3.21782 secs

**THIS PAGE BLANK (USPTO)**